

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: November 8, 2002, 20:31:42 : Search time 1818.72 Seconds
(without alignments)
736.085 Million cell updates/sec

Title: US-09-833-017B-2
Perfect score: 227
Sequence: 1 MKKTLSLKNDKKEIKTDELE.....GSLSTFFFLNRSFTQALCK 46

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -MODEL=frame_plus_p2n_model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09833017/runat_05112002_105348_4834/app_query.fasta_1.398
-DB=GenEmbl -QEXT=fastap -SUFFIX=tde -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017@cgn.1_1616.ernat_05112002_105348_4834 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.hcg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.or.*
21: em.ot.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.to.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.hcg_hum.*
31: em.hcg_inv.*
32: em.hcg_other.*
33: em.hcg_mus.*
34: em.hcg_pln.*
35: em.hcg_rtd.*
36: em.hcg_nam.*
37: em.hcg_vrt.*
38: em.sy.*
39: em.hcg_hum.*
40: em.hcg_mus.*
41: em.hcg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	227	100.0	141	1	AF277152	Streptococ
2	227	100.0	141	1	AF277153	Streptococ
3	227	100.0	141	1	AF277155	Streptococ
4	227	100.0	141	1	AF277156	Streptococ
5	227	100.0	141	1	AF277157	Streptococ
6	220	96.9	141	1	AF277151	Streptococ
7	211	93.0	159	1	AF277154	Streptococ
c 8	69	30.4	8411	6	BD003696	Polynucle
9	69	30.4	10477	1	AE007510	Streptococ
10	69	30.4	10478	1	AE008567	Streptococ
11	69	30.4	232807	2	SPNEU1901	Streptococ
c 12	67	29.5	116557	9	AC107022	Homo sapi
c 13	67	29.5	14935	2	AC009851	Homo sapi
c 14	67	29.5	174028	2	AC067831	Homo sapi
c 15	67	29.5	180389	9	AC063944	Homo sapi
16	65.5	28.9	173868	2	AP001525	Homo sapi
17	65.5	28.9	181083	2	AP001268	Homo sapi
c 18	64.5	28.4	301550	1	AP003192	Clostridi
c 19	64	28.2	118226	9	HS46618	Homo sapi
c 20	63.5	28.0	2455	1	LLJ002203	Lactococc
21	63.5	28.0	10899	1	AF242367	Lactococc
c 22	63.5	28.0	13028	3	AC024836	Caenorhab
c 23	63.5	28.0	41385	3	AF106573	Caenorhab
c 24	63.5	28.0	87246	9	AC002539	Homo sapi
25	63.5	28.0	133925	9	AC004562	Homo sapi
26	63.5	28.0	136421	2	AC006883	Caenorhab
27	63.5	28.0	150789	2	AC006885	Caenorhab
c 28	63.5	28.0	155609	9	HSJ813016	Human DNA
29	62.5	27.5	87863	9	AL513188	Human DNA
30	62.5	27.5	153762	9	AL355338	Human DNA
c 31	62.5	27.5	171031	2	AC104020	Homo sapi
32	62.5	27.5	165967	9	CNS01DVR	Human chr
33	62.5	27.5	213087	2	AC124973	Mus muscu
c 34	62	27.3	110000	2	CEY10568_2	Continuation (3 of
c 35	62	27.3	128746	2	AC094590	Rattus no
36	62	27.3	162427	2	AC115502	Rattus no
37	62	27.3	187260	9	AC018976	Homo sapi
38	62	27.3	277607	3	CEY105E8A	Caenorhab
39	61.5	27.1	157381	9	AL360267	Human DNA
40	61.5	27.1	176085	9	AL391600	Human DNA
41	61.5	27.1	298406	3	CEY7588A	Caenorhab
42	61	26.9	135382	9	AC023798	Homo sapi
43	61	26.9	187732	9	CNS01RHM	Human chr
c 44	61	26.9	229547	2	AC063935	Homo sapi
45	61	26.9	283274	2	AC055759	Homo sapi

ALIGNMENTS

RESULT 1

AF277152
LOCUS Streptococcus mutans strain GB14 DNA linear BCT 07-FEB-2001
DEFINITION precursor (comC) gene, complete cds.

ACCESSION AF277152
VERSION AF277152
KEYWORDS AF277152.1 GI:12698429
SOURCE Streptococcus mutans.
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 141)
AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P., and Cvitkovitch, D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE 21142515
PUBMED 11208787

REFERENCE 2 (bases 1 to 141)
AUTHORS Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada

FEATURES
Location/Qualifiers
1..141
/organism="Streptococcus mutans"
/strain="GB14"
/db_xref="taxon:1309"
1..141
/gene="comC"
1..141
/gene="comC"
/note="comC"
/codon_start=1
/product="competence stimulating protein precursor"
/protein_id="AAK01542.1"
/db_xref="GI:12698430"
/translation="MKKTLKNDKFKETKDELEIIIGSGSLSTFFFLFNRSTQAL GK"
76..138
mat_peptide
76..138
/gene="comC"
/product="competence stimulating protein"

BASE COUNT 58 a 19 c 25 g 39 t
ORIGIN

Alignment Scores:
Pred. No.: 1.01e-25 Length: 141
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-833-017B-2 (1-46) x AF277152 (1-141)

QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGlu 20
|||||
Db 1 ATGAAAAAACACTATCATTAATAAAGAAATTAAGAAATTAAGAACTGATGAATTAGAG 60
|||||

QY 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
|||||
Db 61 ATTATCATTTGGCGGAGCGGAGCGCTATCAACATTTTCCGGCTGTTTAAACAGAGTTT 120
|||||

QY 41 ThrGlnAlaLeuGlyLys 46
|||||
Db 121 ACACAGCTTTGGGAAA 138
|||||

RESULT 2
AF277153
LOCUS Streptococcus mutans strain H7 DNA linear BCT 07-FEB-2001
DEFINITION precursor (comC) gene, complete cds.

AF277153
VERSION AF277153.1 GI:12698431
KEYWORDS Streptococcus mutans.
SOURCE Streptococcus mutans.
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus.

REFERENCE 1 (bases 1 to 141)
AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P., and Cvitkovitch, D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE 21142515
PUBMED 11208787

REFERENCE 2 (bases 1 to 141)
AUTHORS Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada

FEATURES
Location/Qualifiers
1..141
/organism="Streptococcus mutans"
/strain="H7"
/db_xref="taxon:1309"
1..141
/gene="comC"
1..141
/gene="comC"
/note="comC"
/codon_start=1
/transl_table=11
/product="competence stimulating protein precursor"
/protein_id="AAK01543.1"
/db_xref="GI:12698432"
/translation="MKKTLKNDKFKETKDELEIIIGSGSLSTFFFLFNRSTQAL GK"
76..138
mat_peptide
76..138
/gene="comC"
/product="competence stimulating protein"

BASE COUNT 58 a 19 c 25 g 39 t
ORIGIN

Alignment Scores:
Pred. No.: 1.01e-25 Length: 141
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-833-017B-2 (1-46) x AF277153 (1-141)

QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGlu 20
|||||
Db 1 ATGAAAAAACACTATCATTAATAAAGAAATTAAGAAATTAAGAACTGATGAATTAGAG 60
|||||

QY 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
|||||
Db 61 ATTATCATTTGGCGGAGCGGAGCGCTATCAACATTTTCCGGCTGTTTAAACAGAGTTT 120
|||||

QY 41 ThrGlnAlaLeuGlyLys 46
|||||
Db 121 ACACAGCTTTGGGAAA 138
|||||

RESULT 3
AF277155
LOCUS Streptococcus mutans strain UT11 DNA linear BCT 07-FEB-2001
DEFINITION precursor (comC) gene, complete cds.

ACCESSION AF277155
VERSION AF277155.1 GI:12698435
KEYWORDS Streptococcus mutans.

ORGANISM Streptococcus mutans
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 141)
AUTHORS Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE 21142515
PUBMED 11208787
REFERENCE 2 (bases 1 to 141)
AUTHORS Lau,P.C.Y. and Cvitkovitch,D.G.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES source
Location/Qualifiers
1..141
/organism="Streptococcus mutans"
/strain="L711"
/db_xref="taxon:1309"
1..141
/gene="comC"
1..141
/note="comC"
/codon_start=1
/transl_table=1
/product="competence stimulating protein precursor"
/protein_id="AAK01545.1"
/db_xref="GI:12698436"
/translation="MKKTLKNDFKRKDELEIIIGGGSLSTFFRLNRSFTQAL
GK"
mat_peptide 76..138
/gene="comC"
/product="competence stimulating protein"
BASE COUNT 58 a 19 c 25 g 39 t
ORIGIN
Alignment Scores:
Pred. No.: 1,01e-25 Length: 141
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-833-017B-2 (1-46) x AF277155 (1-141)
QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 1 ATGAAAAAACACTATCATTAATAAATGACCTTAAGAAATTAAGACTGATGAATTAGAG 60
QY 21 IleIleIleGlySerGlySerLeuSerThrPheArgLeuPheAsnArgSerPhe 40
Db 61 ATTATCATTTGGCGAAGCGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTTT 120
QY 41 ThrGlnAlaLeuGlyLys 46
Db 121 ACACAAGCTTTGGGAAA 138
RESULT 4
AF277156 141 bp DNA linear BCT 07-FEB-2001
LOCUS Streptococcus mutans strain NG8 competence stimulating protein
DEFINITION precursor (comC) gene, complete cds.
ACCESSION AF277156
VERSION AF277156.1 GI:12698437
KEYWORDS Streptococcus mutans.
SOURCE Streptococcus mutans.
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 141)
AUTHORS Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
J. Bacteriol. 183 (3), 897-908 (2001)

AUTHORS Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE 21142515
PUBMED 11208787
REFERENCE 2 (bases 1 to 141)
AUTHORS Lau,P.C.Y. and Cvitkovitch,D.G.
TITLE Direct Submission
JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES source
Location/Qualifiers
1..141
/organism="Streptococcus mutans"
/strain="NG8"
/db_xref="taxon:1309"
1..141
/gene="comC"
1..141
/note="comC"
/codon_start=1
/transl_table=1
/product="competence stimulating protein precursor"
/protein_id="AAK01546.1"
/db_xref="GI:12698438"
/translation="MKKTLKNDFKRKDELEIIIGGGSLSTFFRLNRSFTQAL
GK"
mat_peptide 76..138
/gene="comC"
/product="competence stimulating protein"
BASE COUNT 58 a 19 c 25 g 39 t
ORIGIN
Alignment Scores:
Pred. No.: 1,01e-25 Length: 141
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-833-017B-2 (1-46) x AF277156 (1-141)
QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 1 ATGAAAAAACACTATCATTAATAAATGACCTTAAGAAATTAAGACTGATGAATTAGAG 60
QY 21 IleIleIleGlySerGlySerLeuSerThrPheArgLeuPheAsnArgSerPhe 40
Db 61 ATTATCATTTGGCGAAGCGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTTT 120
QY 41 ThrGlnAlaLeuGlyLys 46
Db 121 ACACAAGCTTTGGGAAA 138
RESULT 5
AF277157 141 bp DNA linear BCT 07-FEB-2001
LOCUS Streptococcus mutans strain UA159 competence stimulating protein
DEFINITION precursor (comC) gene, complete cds.
ACCESSION AF277157
VERSION AF277157.1 GI:12698439
KEYWORDS Streptococcus mutans.
SOURCE Streptococcus mutans.
ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Streptococcus.
REFERENCE 1 (bases 1 to 141)
AUTHORS Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G.
TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
J. Bacteriol. 183 (3), 897-908 (2001)

Wed Nov 13 10:42:10 2002

TITLE	Direct Submission		
JOURNAL	Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada		
FEATURES	Location/Qualifiers		
source	1. .141		
	/organism="Streptococcus mutans"		
	/strain="BM71"		
	/db_xref="taxon:1309"		
gene	1. .141		
	/gene="comC"		
CDS	1. .141		
	/gene="comC"		
	/note="ComC"		
	/codon_start=1		
	/transl_table=11		
	/product="competence stimulating protein precursor"		
	/protein_id="AAK01541.1"		
	/db_xref="GI:12698428"		
	/translation="MKKTPSLKNDKFEIKTDELIIGSGSLSTFFRFLNRSFTQAL GK"		
mat_peptide	76. .138		
	/gene="comC"		
BASE COUNT	58 a 20 c 25 g 38 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1.14e-24	Length:	141
Score:	220.00	Matches:	45
Percent Similarity:	97.83%	Conservative:	0
Best Local Similarity:	97.83%	Mismatches:	1
Query Match:	96.92%	Indels:	0
DB:	1	Gaps:	0
US-09-833-017B-2 (1-46) x AF277151 (1-141)			
Qy	1	MetLysLysThrLeuSerLeuLysAsnAspPheLysGluLeLysThrAspGluLeuGlu	20
Db	1	ATGAAAAAACACCATCATATAAAATGACTTAAAGAAATTAAGACTGATGAATTAGAG	60
Qy	21	lletllelGlyGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe	40
Db	61	ATTATCATTTGGCGGAAGCGAGCCATACACATTTTCCGGCTCTTTAACACAGAAGTTT	120
Qy	41	ThrGlnAlaLeuGlyLys	46
Db	121	ACACAAGCTTTGGAAAA	138
RESULT 7			
AF277154	159 bp DNA linear BCT 07-FEB-2000		
LOCUS	Streptococcus mutans strain JHI005 competence stimulating protein		
DEFINITION	precursor (comC) gene, complete cds.		
ACCESSION	AF277154		
VERSION	AF277154.1 GI:12698433		
KEYWORDS	Streptococcus mutans.		
SOURCE	Streptococcus mutans		
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
REFERENCE	1 (bases 1 to 159)		
AUTHORS	Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G.		
TITLE	Natural genetic transformation of Streptococcus mutans growing in biofilms		
JOURNAL	J. Bacteriol. 183 (3), 897-908 (2001)		
MEDLINE	21142515		
PUBMED	11208787		
REFERENCE	2 (bases 1 to 159)		
AUTHORS	Lau,P.C.Y. and Cvitkovitch,D.G.		
TITLE	Direct Submission		
JOURNAL	Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada		
FEATURES	Location/Qualifiers		

MEDLINE	21142515
PUBMED	11208787
REFERENCE	2 (bases 1 to 141)
AUTHORS	Lau,P.C.Y., and Cvitkovitch,D.G.
TITLE	Direct Submission
JOURNAL	Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES	Location/Qualifiers
source	1..141 /organism="Streptococcus mutans" /strain="UA159" /db_xref="taxon:1309"
gene	1..141 /gene="comC"
CDS	1..141 /gene="comC" /note="ComC" /codon_start=1 /transl_table=11 /product="competence stimulating protein precursor" /protein_id="AAK01547.1" /db_xref="GI:12698440" /translation="MKKTSLKNDFKEIKTDELEIIIGSGSLTFRLNRSFTQAL GK" 76..138 mat_peptide /gene="comC" /product="competence stimulating protein"
BASE COUNT	58 A 19 C 25 G 39 T
ORIGIN	
Alignment Scores:	
Pred. No.:	1 01e-25 Length: 141
Score:	227.00 Matches: 46
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	100.00% Indels: 0
DB:	1 Gaps: 0
US-09-833-017B-2 (1-46) x AF277157 (1-141)	
Oy	1 MetLysLysThrLeuSerLeuLySAsnAaspHelysGlufleLysThrAspCluLeuGlu 20 Db 1 ATGAATAAACACATCATCTATTAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAGAG 60 Oy 21 llelleileGLyGSerGLySerLeuSerThrPhePheArGLeuPheAsnArgSerPhe 40 Db 61 ATTATCATGGCGGNAGCGGAGCCATCAACATTTTTTCGGCTGTTTACACAAGGTTTT 120 Oy 41 ThrGlnAlaLeuGLyLys 46 Db 121 ACACAAGCTTTGGGAAA 138
RESULT 6	
LOCUS	AF277151 141 bp DNA linear BCT 07-FEB-2001
DEFINITION	Streptococcus mutans strain BM71 competence stimulating protein precursor (comC) gene, complete cds.
ACCESSION	AF277151
VERSION	AF277151.1 GI:12698427
KEYWORDS	Streptococcus mutans.
SOURCE	Streptococcus mutans
ORGANISM	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Streptococcus. 1 (bases 1 to 141) Li,Y.H., Lau,P.C., Lee,J.H., Ellen,R.P. and Cvitkovitch,D.G. Natural genetic transformation of Streptococcus mutans growing in biofilms
REFERENCE	J. Bacteriol. 183 (3), 897-908 (2001)
AUTHORS	21142515
TITLE	11208787
JOURNAL	2 (bases 1 to 141)
MEDLINE	Lau,P.C.Y., and Cvitkovitch,D.G.
PUBMED	
REFERENCE	

FEAT	FT source	Location/Qualifiers 1..8411 /organism='Unidentified'
FEATURES	source	Location/Qualifiers 1..8411 /organism='unidentified' /db_xref='taxon:32644'
BASE COUNT	2741 a 1386 c 1639 g 2645 t	
ORIGIN		
 Alignment Scores:		
Score:	8.68	Length: 8411
Percent Similarity:	69.00	Matches: 19
Best Local Similarity:	50.88%	Conservative: 10
Query Match:	33.33%	Mismatches: 16
	30.40%	Indels: 12
DB:	6	Gaps: 2
 US-09-833-017B-2 (1-46) x BD003696 (1-8411)		
QY	2 LysLysThrLeuSerLeuLysAsnAspPhe-	-----LysGlu 13
Ddb	3566 AAAAGACGTTTACAACTGATAATGTCTTCCTCCTTTAATCTAGGATATCAATAAGAG	3507
QY	14 IleLysThrAspGluLeuGluIleIleGlySerGlySerLeuSer-	-----30
Ddb	3506 AAAAACAATCTTCAATATCAATGATTAGTTCCCTGGAAATTCACAGTAGTATCATTTGAATCT	3447
QY	31 ---ThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys	46
Ddb	3446 CCTACTTCTTCTGCTTTAGATACAAAACATTTCCACCGAGGCAATCAAAAA	3396
RESULT 9		
LOCUS	AE007510	10477 bp DNA linear BCT 31-AUG-2001
DEFINITION	Streptococcus pneumoniae TIGR4 section 193 of 194 of the complete genome.	
ACCESSION	AE007510	AE005672
VERSION	AE007510.1	GI:14973736
KEYWORDS		
SOURCE	Streptococcus pneumoniae TIGR4	
ORGANISM	Streptococcus pneumoniae TIGR4	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
REFERENCE		Streptococcus
AUTHORS	1 (bases 1 to 10477)	Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Salzberg,S.L., Lewis,M.R., Radune,D., Holtzapple,E., Khouri,H., Wolf,A.M., Utterback,T.R., Hansen,C.L., McDonald,L.A., Feldblyum,T.V., Angiuoli,S., Dickinson,T., Hickey,E.K., Holt,I.E., Loftus,B.J., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.
TITLE	Complete genome sequence of a virulent isolate of Streptococcus pneumoniae	
JOURNAL	Science 293 (5529), 498-506 (2001)	
MEDLINE	21357209	
PUBMED	11463196	
REFERENCE	2 (bases 1 to 10477)	
AUTHORS	Tettelin,H., Nelson,K.E., Paulsen,I.T., Eisen,J.A., Read,T.D., Peterson,S., Heidelberg,J., Deboy,R.T., Haft,D.H., Dodson,R.J., Durkin,A.S., Gwinn,M., Kolonay,J.F., Nelson,W.C., Peterson,J.D., Umayam,L.A., White,O., Lewis,M.R., Radune,D., Holtzapple,E., Khouri,H., Wolf,A.M., Utterback,T.R., Hansen,C.L., McDonald,L.A., Feldblyum,T.V., Angiuoli,S., Gesuwan,P., Hickey,E.K., Holt,I.E., Loftus,B.J., Ujwal,M.L., Yang,F., Smith,H.O., Venter,J.C., Dougherty,B.A., Morrison,D.A., Hollingshead,S.K. and Fraser,C.M.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-JUN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA	
FEATURES	Location/Qualifiers	
source	1..10477	

```

/organism="Streptococcus pneumoniae TIGR4"
/strain="TIGR4"
/db_xref="taxon:170187"
110..478
/gene="SP2226"
CDS
110..478
/gene="SP2226"
/feature="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK76274.1"
/db_xref="GI:1497373"
/translation="MEYKLFEEFTTLOALFKELGITHSGGAIKSFSEHSVYFNRELE
SRGKALRIGKVDIPDNNDILLITPTSEQDEYQADKVEKRIAKLVKMKMGKVK
DKSKPASPSQQAQPPFGR"
481..1578
/gene="SP2227"
CDS
481..1578
/gene="SP2227"
/codon_start=1
/transl_table=11
/product="recf protein"
/protein_id="AAK76275.1"
/db_xref="GI:14973738"
/translation="MWLQHLKTRFYKFKIDPNKLVFLGRNAQGTNMLEAIY
FLATRSRTRTDKMLIHFEDEQLHLGLVOKKTSIPLEITLTKGRVTKVNLKQA
RLSDVGHMNVLPAPEDQLIKGAPSTRKFRIDMELGQIKPTVLSLDTNHLKOR
NYLAKAQIDTFSLVDLDOLVQGCVMHRLDFIKKLESGRKHFKFELSNQIEL
SISYQSSNIDKQNLSPKIALESRSRLFKNTGVGPHRDDISFYINGMDASFG
SQGRKSLQNLKAEIEMESITTESPILLDLDVNSLONTRKLLKLETISQITF
ITTSLDHLQNPENLSITODGKAAYNGN"
complement(1629..3107)
/gene="SP2228"
CDS
complement(1629..3107)
/gene="SP2228"
/feature="identified by match to PFAM protein family HMW
pf00478"
/codon_start=1
/transl_table=11
/product="inosine-5'-monophosphate dehydrogenase"
/protein_id="AAK76276.1"
/db_xref="GI:14973739"
/translation="MSNWDTKFLKGFEDVDVLLIPAESVLPNDADLTTLKADNLT
NPIITAMADVTEQMAIATARAGGLGVHKNMSIAQQADEVRKYKRSNGVIIDPF
FUTPEHTIAEDELGRIRISGVPPVTEFKRLVGLTNRDLRFSDINQPSINHMT
SENLTAPVGTDLTAESILQEHREKLPVDEEGLSGLTINDIEKVEFFNAARD
EGRLLVAGVGVTSDFEAEALFEAGADAIVIDTAHGSAGVLRKIAETRAHFPDR
TLIAGNIAEAGARALYBAGVDVVKVIGPGSICITRVIAGVGPVQTAIYDAAVAR
EYGTILADGGIKYSGDIVKALAAGGNVLMGSMFAGTDEAPGETEIFQGRKFTYRG
MGSIAAMKGGSDRYFGVSNANKLVPEGTGRVAYKGAADIVFQMIGGIRSGMGY
CGAANKLHONAOPIEMSGAGLKESHDPDVQITNEAPNYSM"
complement(3259..4284)
/gene="SP2229"
CDS
complement(3259..4284)
/gene="SP2229"
/feature="identified by match to TIGR protein family HMW
TIGR00233"
/codon_start=1
/transl_table=11
/product="tryptophanyl-tRNA synthetase"
/protein_id="AAK76277.1"
/db_xref="GI:14973740"
/translation="MTKPIILTDGRPTGKLHIGYVGLSKNRVLLQBEDKYDMFVFLA
DQALTDHAKDPTTEISIGNVALDYLAVGLDPNKSIFIQSQIPELAELSMYIMLV
SIALRENPRTQTEISOKFGESITPTGLVPIAQAADITAFKANYVPVGTDOKMPIE
QRTREIVSFNNAYNDVLEPEGIVPENRAGRLPGLDGNKSKSLNNGIYLADAD
TLRKVMNWTDPHTRVDEPKLEGNMVFXLDVFGRPDQAEIADMKRYQRGIG
DYKTRYLLEILRELGPERRIEFAKMDGEVINMLQKGSERAREVAGOTLSEVGA
MGLHFN"
4490..6112
/gene="SP2230"
CDS
4490..6112

```

```

/gene="SP2230"
/codon_start=1
/transl_table=11
/product="ABC transporter, ATP-binding protein"
/protein_id="AAK76278.1"
/db_xref="GI:14973741"
/translation="MLTVSDVSLRFSRDKLEDDVNIKFTSGNTYGLIGANGACKSTFL
KILAGDIEPTGHSILGPDERKLSVLRNHEDIYERADIVVIMNKKLSIMKEKDAI
YMKEDFDEDCVRAAELEGEPAELGWEASEASQLQNLNIPELHYQMSLANGE
KVYLLAKALFKGKPDVLLIDPTNGLDISITWIEDFLIDFNTVIVSHDRFLNKKV
CTHMADLDFGKIKLYGVNYDFKESSELAALKLADNAKAEKIKQLQEFARFSANA
SKRSQATSEKMKLDKILEIEIVPSRRKPIFNKAREIGNDILLIVENTVTKIDGETI
LONTSFILRPDKTALIGONDIQTALIRAIMGDIYEGVVKWGVTSQSYLPKDNKA
DPAGESILMLRQFASKEEDNTFLRGLRMLFSGDEVNKPVNLVSGEKVVRMLS
KMLKNSVILVDDPTNHDLESISSLDGLKNFKESILIFASHDHEFIQTLANHIIVL
SKNGVIDRIDETDEFELENAEQAKVXELMKRD"
6174..8726
/gene="SP2231"
CDS
6174..8726
/gene="SP2231"
/codon_start=1
/transl_table=11
/product="ABC transporter, permease protein, putative"
/protein_id="AAK76279.1"
/db_xref="GI:14973742"
/translation="MKGFLKTYRTYFISFIPVIMSGVLSQSIYWNDSNPLGSG
PHQVYIPDALNRNLHGNSNLYFTTSGLGNFYALSSYVGLSFLAPLVYFDTLNM
DAIYLLTGKGLIGLSTFSLNKLQFOSIPQILKALSTSYALMSTVSOLEIKTWLD
VPIILITLGLHLLITEKLLLYFTLSILFIONYEGVWVFLIFLFWLQLSWDF
KTRKSSVLDIFSVISLAGMASLMTLFTLDQTHGKLEVEYKFKOTESWILDLFAK
QFIPQGTGKAPIMFVGLPFLFTLFTLTKSKFKHVKLIYIFFAFLIASFVIE
ALDFWQGMHTNMFLHRYAMFSTLLIYAAEVLKRLKELKYNMFLSVLFAVGLF
ATYIKSHYSFELTDNLITLLEFVYVSLLLAVIKKIFSVNLFALISLFLVENS
NASOMDGIKEMGFASRSAYSRDIPAMESFTYIGNQFTRTEKLTQGTGNDKMFNY
NGISOFSSVRNRSSTLDLKGKSGTNLNNYANNISLADSLFGIQYNISDSIDK
YGFIDYKDNLTLYNQYSLPIAVASQSYNDVKNETHLDNQASFQNLQANVDFY
PSPIPYKTEKIENTNDLISVSSNEDALQVLEVPENSQVYLSFINLHFSNDQK
KVDILVNGEKKTFITDVFNSFNLCYTKKKTENINVSFPNGSQVSFEPTFRLDTK
TPTLAIQIKIEQPVVTSKNKVFATYDQVDSITPTIPYDKGWSAQDQKAKIEIK
AQTFMKYDIPKGGKITLSPFNGFITGATCSFTSLLEFGIYNHRRKSSKA"
complement(8827..9506)
/gene="SP2232"
/feature="this region contains an authentic frame shift and
is not the result of a sequencing artifact; identified by
Glimmer2; putative; conserved hypothetical protein,
authentic frameshift"
/complement(9591..10358)
/gene="SP2233"
CDS
complement(9591..10358)
/gene="SP2233"
/feature="identified by Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAK76280.1"
/db_xref="GI:14973743"
/translation="MIGISLIPDLNIFLSSMMDPYGQLSDLPVAVVNNKREASNG
NTMAGKDMVSNLKENKTLDFHFDVEEGKGLGDDGYVMVTLPSDLSEKTTLSNI
QSTAAYQSLTSEQOTEISDSVONSSTIOSAOSIVALVODLOGSLNQLQSSNLS
LKNSQNSPITSTSLIGLSGLTEIOGDTSKLVPASQSIASGVNAYTGVQKVSOG
ASQSEKNATLTLGSKLVSGSNLTQKSSRLTAGVG"
BASE COUNT 3366 a 2047 c 1699 g 3365 t
ORIGIN
Alignment Scores:
Pred. No.: 11.3 Length: 10477
Score: 69.00 Matches: 19
Percent Similarity: 50.88% Conservativeness: 10
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 30.40% Indels: 12
DB: 1 Gaps: 2
US-09-833-017B-2 (1-46) x AE007510 (1-10477)

```


SKSROATSKRMKLEIEETVPSSRRKYPINFAKREIGNDLLTVENITVKIDGETI
 LNI5FILPPDKTALIGNDIOITRALIRMGIDIDYEGTVKMGVTTOSYLPKNSA
 DRAGESILDLWROFASKEEDNTFLRGLRMLFSGDEVNKPVNLVSGEKVRVLS
 KMLKLSNVLDLDFTHLDESISLNDGLKNFKESLIIFASHDEFIOTLANHIIVL
 SRNGVIDRIDETVDEFLNAEVQAKVRELWKD"
 6174. .8726
 /gene="spr2036"
 6174. .8726
 /gene="spr2036"
 /codon_start=1
 /transl_table=11
 /label=spr2036
 /product="Conserved hypothetical protein"
 /protein_id="AAL00838.1"
 /db_xref="GI:15459743"

gene

CDS

/translation="MKSFKTYRTYFTSLIIPVIMSGVVLVSGIYVNSDNPGLGDD
 PQVYIFVALRNLTHGNSLFTFTSGLGNFVALSSYLGSLAPLVYFFDLNMP
 DAVYLTLLKGLLGSNFFSLKLFQSIPTLKLALSTYALMSFTVSOLEIKTWLD
 VPIILLITGLHLLITEKKLLLYFTSLIFQNYFYGMVVLFIWYLQISWDF
 QTKGSVDYFIVISFLAGMASLIMTFLFDLOTHGKLEFVTKFOTESSWLDLFAK
 QTKGSVDYFIVISFLAGMASLIMTFLFDLOTHGKLEFVTKFOTESSWLDLFAK
 ADLEWQGMHTPNMFLHRYAWIFSTLIYTAAEVIRKELKVMNPLVSLFVAGFL
 ATYILKSHYSLTDNLITLLEFLVYVYLLAVIRKELKVMNPLVSLFVAGFL
 NASSQDGLKEWGSFASRSYSDILAMESFTYIGNQFTREKLTQTCNDSMKFN
 YGFIQDKDNLTYENQYSLPIAVASQSYNDVKNHTLNOQSLFNLANVNFQK
 FSPIDYKTEKIENTNOLISVTSNEDDAIQLEVPNSQVYLSFLNHFNSKQK
 KYDVLNKGKIFTIDNVSFNLTKRKTNINVSFGNSQVSFESPTFIRLDIK
 TATEAQIKQPVTVSTKNKVFATYDVOODTSIFPTIPYDKGWSAYODGKKIEIKQ
 AQGTGMKVIDIPKGGTITLSPFINGFITGACISFTSLLLFGIYNHRRKSSKA"
 complement(8827. .9141)
 /gene="spr2037"
 complement(8827. .9141)
 /gene="spr2037"
 /codon_start=1
 /transl_table=11
 /label=spr2037
 /product="Conserved hypothetical protein, truncation"
 /protein_id="AAL00839.1"
 /db_xref="GI:15459744"

gene

CDS

/translation="MAGILVYGVQLIGLITANHEMRIFILITSLVFSMVTTLATW
 NSRIGAFSLILLQLASSAGTYPLATNDPFRINPLPNSYSVGLRQISINKS
 FS"
 complement(9169. .9507)
 /gene="spr2038"
 complement(9169. .9507)
 /gene="spr2038"
 /codon_start=1
 /transl_table=11
 /label=spr2038
 /product="Conserved hypothetical protein, truncation"
 /protein_id="AAL00840.1"
 /db_xref="GI:15459745"

gene

CDS

/translation="MADGSRKLAEGGKLTSGLEDLTQGLASLGGLGNASDLKSVS
 TESKNAEILSNPLNLTNDQVPVNGIATPYMISVALFFAAISTNMFIAKLPSGRH
 PESRWAWLKS"
 complement(9592. .10359)
 /gene="spr2039"
 complement(9592. .10359)
 /gene="spr2039"
 /codon_start=1
 /transl_table=11
 /label=spr2039
 /product="Conserved hypothetical protein, truncation"
 /protein_id="AAL00841.1"
 /db_xref="GI:15459746"
 /translation="MIGISLIPDLNIIIFLSSWWDYQGLSDLPVAVVNDKEASYNG
 NMAIGKDMVSNLKNKTLDFHVEEGKRGLEDGYIMVTLPPSLSEKTTLSNI
 QSTAAVQLTSEQTETSDSVSQNSQNTSDSIQSAQISVALVDQLQGSLENLQNSNLSL

gene

CDS

Alignment Scores:
 Pred. No.: 11.3 Length: 10478
 Score: 69.00 Matches: 19

Percent Similarity: 50.88% Conservative: 10
 Best Local Similarity: 33.33% Mismatches: 16
 Query Match: 30.40% Indels: 12
 DB: 2 Gaps: 2

US-09-833-017b-2 (1-46) x AE008567 (1-10478)

Qy 2 LysLysThrLeuSerLeuLysAsnAspPhe-----LysGlu 13
 ||||| |||||
 Db 8247 AAAAGACTTTTACAACTGATAATGCTCTCTCTTTAACTAGGATATACTAAAGAG 8306
 ||||| |||||
 Qy 14 IleLysThrAspGluLeuGluIleIleIleGlyGlySerGlySerLeuSer-----30
 ||||| |||||
 Db 8307 AAAAATACTTCAATCAATAGTCTAGTTCTCCCTGGAAATTCACAAAGTATCATTTGAATCT 8366
 ||||| |||||

Qy 31 ---ThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 46
 ||||| |||||
 Db 8367 CCTACCTCTCATGCTTTAGATACCAAACTTTCACCGAGCAATCAAAA 8417
 ||||| |||||

RESULT 11

SPNEU1901 232807 bp DNA linear HTG 12-JUL-2001
 Streptococcus pneumoniae clone G34, *** SEQUENCING IN PROGRESS ***
 DEFINITION in ordered pieces

ACCESSION AL449923

VERSION AL449923.1 GI:11545148

KEYWORDS HTG; HTGS_PHASE2.

SOURCE Streptococcus pneumoniae.

ORGANISM Streptococcus pneumoniae

Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

Streptococcus.

REFERENCE 1 (bases 1 to 232807)

AUTHORS
 Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,
 Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de
 Francesco,M., Polissi,A., Buell,G., Feger,G., Garcia,E., Peitsch,M.

TITLE Annotated draft genomic sequence from a Streptococcus pneumoniae

JOURNAL type 19F clinical isolate

MEDLINE Microb. Drug Resist. 7 (2), 99-125 (2001)

PUBMED 21335329

REFERENCE 11442348

AUTHORS 2 (bases 1 to 232807)

Dopazo,J., Mendoza,A., Herrero,J., Caldara,F., Polissi,A.,
 Humbert,Y., Friedli,L., Guerrier,M., Grand-Schenk,E., Gandin,C., de
 Francesco,M., Buell,G., Feger,G., Garcia,E., Peitsch,M. and
 Garcia-Bustos,J.F.

TITLE Direct Submission

JOURNAL Submitted (31-OCT-2000)

COMMENT Severe Ochoa 2, 28760 Tres Cantos, SPAIN

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

FEATURES

source

1. .232807

/organism="Streptococcus pneumoniae"

/serotype="19F"

/db_xref="taxon:1313"

/clone="G54"

BASE COUNT 70984 a 43538 c 47937 g 70253 t 95 others

ORIGIN

Alignment Scores:

Pred. No.: 461 Length: 232807

Score: 69.00 Matches: 19

Percent Similarity: 50.88% Conservative: 10

Best Local Similarity: 33.33% Mismatches: 16

Query Match: 30.40% Indels: 12

DB: 2 Gaps: 2

US-09-833-017b-2 (1-46) x SPNEU1901 (1-232807)

Qy 2 LysLysThrLeuSerLeuLysAsnAspPhe-----LysGlu 13

```

|||||...|||
Db 48135 AAAAGACTTTTACAACTGATAGTCCTCTCTCTTAACTAGGATATACATAAGAG 48194
|||||
Qy 14 IlelysthrspgluLeuGlulIlelleGlySerGlySerLeuSer----- 30
|||||
Db 48195 AAAAATACTTTCATCAATGTTAGTTCCTCGGAAATTCACAAGTATCATGAATCT 48254
|||||
Qy 31 ---PhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 46
|||||
Db 48255 CCTACCTTCATCGTTAGATACAAACATTCACCGAGGCAATTCACAAA 48305
|||||
RESULT 12
AC107022/c
LOCUS
DEFINITION Homo sapiens 3 BAC RP11-417H23 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC107022
VERSION AC107022.4
KEYWORDS GI:20334515
SOURCE HTG.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 116557)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayalew,M., Banks,T.,
Barbala,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denna,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,R.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
Hollins,B., Homsif,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudas,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulsegh,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenwo,S., Ogih,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E.,
Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wleczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Zyltor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 116557)
Worley,K.C.
Direct Submission
Submitted (14-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

```

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 116557)

Worley,K.C.

Direct Submission

Submitted (22-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USAREFERENCE
AUTHORS
TITLE
JOURNAL

4 (bases 1 to 116557)

Worley,K.C.

Direct Submission

Submitted (27-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Apr 27, 2002 this sequence version replaced gi:18449824.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1..116557 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="3" /clone="RP11-417H23"
misc_feature	1..1995 /note="overlaps bases 190327..192322 of clone AC074043"
repeat_region	/function="clone overlap" complement(1..77) /rpt_family="THE1C"
repeat_region	170..561 /rpt_family="LIM4" complement(733..956) /rpt_family="LTR40b"
repeat_region	957..1068 /rpt_family="MER81" complement(1069..1363) /rpt_family="LTR40a"
STS	1518..1651 /standard_name="75005"
repeat_region	1760..1787

[illegible]

All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www.seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: 12180
 Center clone name: 417_H23
 ----- Summary Statistics

Sequencing vector: M13; M77815; 98% of reads
 Sequencing vector: Plasmid; n/a; %0.f% of reads
 2.466995599408Chemistry: Dye-primer-amersham; 4% of reads
 Chemistry: Dye-terminator Big Dye; 96% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 143822 bases at least Q40
 Consensus quality: 146616 bases at least Q30
 Consensus quality: 147852 bases at least Q20
 Insert size: 182000; agarose-fp
 Insert size: 148935; sum-of-contigs
 Quality covera.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 13133: contig of 13133 bp in length
 * 13134 13233: gap of 100 bp
 * 13234 18557: contig of 5324 bp in length
 * 18558 18657: gap of 100 bp
 * 18658 35242: contig of 16585 bp in length
 * 35243 35342: gap of 100 bp
 * 35343 77529: contig of 42187 bp in length
 * 77530 77629: gap of 100 bp
 * 77630 149335: contig of 71706 bp in length.

FEATURES

source
 1. 149335
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosomes="3"
 /map="3"
 /clone="Rp11-417H23"
 /clone_lib="RPC1-11 Human Male BAC"
 misc_feature
 1. 13133
 /note="assembly_fragment
 clone_end:SP6
 vector_side:left"
 misc_feature
 13234..18557
 /note="assembly_fragment"
 18658..35242
 /note="assembly_fragment"
 35343..77529
 /note="assembly_fragment"
 77630..149335
 /note="assembly_fragment
 clone_end:T7
 vector_side:right"
 BASE COUNT 46161 a 28722 c 28526 g 45525 t 401 others
 ORIGIN

Alignment Scores:
 Pred. No.: 542 Length: 149335
 Score: 67.00 Matches: 15
 Percent Similarity: 62.86% Conservative: 7
 Best Local Similarity: 42.86% Mismatches: 13
 Query Watch: 29.52% Indels: 0
 DB: 2 Gaps: 0

US-09-833-017b-2 (1-46) x AC009651 (1-149335)

QY 2 LysLysThrLeuSerLeuLysAsnAspPhelYsGluLeuLysThrAspGluLeuGluLeu 21
 ||||| : : : ||| : : : ||||| : : : ||||| : : : ||||| : : :
 Db 35536 AAAAAACACATCAACTCTCAAGAGATTTTGAAGAAAGGGAACAGCAATGAGTG 35477
 ||| : : : ||| : : : ||||| : : : ||||| : : : ||||| : : :
 QY 22 IleIleGlyGlySerLeuSerThrPhePheArgLeuPhe 36
 ||| : : : ||| : : : ||||| : : : ||||| : : : ||||| : : :
 Db 35476 ACTGGAGGGAATAATGGGTCATCATCAATTTTGTGTTT 35432

RESULT 14

AC067831/c

LOCUS

DEFINITION

AC067831

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

Barbaria, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Cahan, R.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, D., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louissegh, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tasey, J., Taylor, C., Taylor, I., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, L., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 180389)
Worley, K.C.

Direct Submission
Submitted (22-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 180389)
Worley, K.C.

Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 180389)
Worley, K.C.

Direct Submission
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 28, 2002 this sequence version replaced gi:19718609.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and

local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT

FEATURES	Source	Location/Qualifiers
		1. .180389
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clone="RP11-446H18"
misc_feature		1. .2015
		/note="overlaps bases 190308..192322 of clone AC074043"
repeat_region		/function="clone overlap"
		complement(1..96)
repeat_region		/rpt_family="THE1C"
		189..580
repeat_region		/rpt_family="L1M4"
		complement(752..975)
repeat_region		/rpt_family="L1R40b"
		976..1087
repeat_region		/rpt_family="MER81"
		complement(1088..1382)
repeat_region		/rpt_family="L1R40a"
		1537..1670
STS		/standard_name="75005"
repeat_region		1779..1807
repeat_region		/rpt_family="(A)n"
		complement(2187..2700)
repeat_region		/rpt_family="MLT1E2"
		4696..4823
repeat_region		/rpt_family="M1R"
		4906..5055
repeat_region		/rpt_family="MLT1J2"
		5081..5265
repeat_region		/rpt_family="MLT1J1"
		6184..6204
repeat_region		/rpt_family="AT_rich"
		6675..6706
repeat_region		/rpt_family="(TTG)n"
		6707..7040
repeat_region		/rpt_family="MLT1J2"
		7531..7845
repeat_region		/rpt_family="L1R16B"
		7926..7960
repeat_region		/rpt_family="AT_rich"
		8015..8041
repeat_region		/rpt_family="(TTG)n"
		complement(8042..8333)
repeat_region		/rpt_family="AluJo"
		complement(8538..8867)
repeat_region		/rpt_family="MLT1E"
		complement(8994..9058)

```

repeat_region /rpt_family="MLT1F"
9060. .9085
/rpt_family="AT_rich"
repeat_region 9203. .9237
/rpt_family="AT_rich"
STS 9335. .9525
/standard_name="8620"
9431. .9488
/rpt_family="(TAGA)n"
repeat_region 10847. .10885
/rpt_family="(CA)n"
repeat_region complement(10888..11004)
/rpt_family="MLT1J2"
repeat_region complement(11035..11143)
/rpt_family="MLT1J2"
repeat_region complement(11196..11508)
/rpt_family="AluJb"
11554. .11640
/rpt_family="Charlie4"
repeat_region 11831. .12154
/rpt_family="LTRL6C"
repeat_region 13419. .13451
/rpt_family="(CAAAA)n"
repeat_region complement(13731..13806)
/rpt_family="MLT1G"
repeat_region complement(13911..14121)
/rpt_family="MLT1G1"
14122. .14427
/rpt_family="AluJb"
repeat_region complement(14428..14466)
/rpt_family="MLT1G1"
14641. .15655
/rpt_family="L1M4"
repeat_region 15738. .16256
/rpt_family="L1M4"
repeat_region 16266. .16564
/rpt_family="AluSx"
repeat_region 16593. .18442
/rpt_family="L1MEL"
18468. .18770
/rpt_family="AluJb"

```

```

Alignment Scores:
Pred. No.: 680 Length: 180389
Score: 67.00 Matches: 15
Percent Similarity: 62.86% Conservative: 7
Best Local Similarity: 42.86% Mismatches: 13
Query Match: 29.52% Indels: 0
DB: 9 Gaps: 0

```

```

US-09-833-017B-2 (1-46) x AC063944 (1-180389)
QY 2 LysLysThrLeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGluLe 21
DB 61451 AAAAAACACATCACTCAAGAGCTTTTGAGAGAGGGAACAGACGAGTGGTG 61392
QY 22 IleIleGlyGlySerGlySerLeuSerThrPhePheArgLeuPhe 36
DB 61391 ACTGGGGGAAAAATGGGGTCATCATCATTTTGTGTTTGT 61347

```

```

Search completed: November 8, 2002, 22:56:29
Job time : 1885.72 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 20:29:27 ; Search time 229.313 seconds
(without alignments)
451.748 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 227

Sequence: 1 MKKTLKNDPKFKTDELE.....GSLSTFRLNRSFTQALGK 46

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09833017/runat_05112002_105348_4824/app_query.fasta_1.398
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017.CGN1.1.125 @runat_05112002_105348_4824 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	227	100.0	141	24	Streptococcus muta
2	227	100.0	141	24	Streptococcus muta
3	227	100.0	2557	24	Streptococcus muta
4	227	100.0	2557	24	Streptococcus muta
5	104	45.8	63	24	Streptococcus muta
6	104	45.8	63	24	Streptococcus muta
c 7	69	30.4	8411	19	Streptococcus pneu
c 8	59	26.0	681	24	Oligonucleotide fo
c 9	59	26.0	681	24	Oligonucleotide fo
c 10	59	26.0	1127	22	Peppermint plant o
c 11	59	26.0	1481	22	Nucleotide sequenc
12	58.5	25.8	416	24	Bacillus clausii g
13	58.5	25.8	885	24	Oligonucleotide fo
c 14	58.5	25.8	885	24	Oligonucleotide fo
c 15	58.5	25.8	1530	24	Streptococcus poly
c 16	58.5	25.8	1533	22	Streptococcus pyog
c 17	58.5	25.8	37661	22	Human neurokinin B
c 18	58	25.6	691	18	DNA encoding a Sta
c 19	58	25.6	4702	18	Staphylococcus aur
c 20	58	25.6	6389	24	Human immune syste
c 21	57	25.1	942	24	Bacillus lichenifo
c 22	57	25.1	48667	24	Human transporter
c 23	56.5	24.9	843	24	Bacillus lichenifo
c 24	56.5	24.9	1061	11	Glyceraldehyde-pho
c 25	56.5	24.9	13884	20	Enterococcus faeca
26	56	24.7	2820	23	Enterococcus faeca
27	56	24.7	5480	20	Enterococcus faeca
28	56	24.7	8467	24	Human chemically m
c 29	56	24.7	465237	24	Human oestrogen re
c 30	56	24.7	465237	24	Human oestrogen re
31	55.5	24.4	1251	24	Listeria monocytog
32	55.5	24.4	1662	17	Newcastle disease
33	55.5	24.4	3825	8	Sequence of Newcas
34	55.5	24.4	4177	15	SfiI fragment cont
35	55.5	24.4	4177	15	SfiI fragment cont
36	55.5	24.4	4177	18	SfiI fragment enco
37	55.5	24.4	4177	20	Seq ID No: 12 of 0
38	55.5	24.4	4177	21	SfiI fragment comp
39	55.5	24.4	4177	21	cdNA encoding NDV
40	55.5	24.4	15186	21	Newcastle disease
41	55	24.2	367	22	Human polynucleoti
42	55	24.2	1752	22	S. epidermidis ope
43	55	24.2	1824	24	Fungal ZBC gene se
44	55	24.2	3033	24	Staphylococcus epi
45	55	24.2	3055	22	S. epidermidis gen

ALIGNMENTS

RESULT 1
AAD32791
ID AAD32791 standard; DNA; 141 BP.
XX
AC AAD32791;
XX
DT 01-JUL-2002 (first entry)
DE
XX Streptococcus mutans comC gene.

Competence signal peptide; CSP; histidine kinase; HK; therapy; caries;
KW endocarditis; immunopurification; antibacterial; antiinflammatory;
KW genetic competence assay; vaccine; gene; comC gene; ds.
OS Streptococcus mutans.

XX Key Location/Qualifiers
FH 1..141
FT CDS

```
FT      /*tag= a
FT      /product= "S. mutans ComC protein"
XX
XX PN      CA2302861-A1.
XX
XX PD      10-OCT-2001.
XX
XX PF      10-APR-2000; 2000CA-2302861.
XX
XX PR      10-APR-2000; 2000CA-2302861.
XX
XX PA      (LAUP/) LAU P C Y.
XX PA      (CVIT/) CVITKOVITCH D G.
XX PA      (LIYH/) LI Y H.
XX
XX PI      Lau PCY, Cvitkovitch DG, Li YH;
XX
XX PD      WPI: 2002-242151/30.
XX
XX DR      P-PSDB; AAE20494.
XX
XX PT      Novel compound that inhibits binding of competence signal peptide of
XX PT      Streptococcus mutans to S. mutans histidine kinase, useful for treating
XX PT      or preventing caries or endocarditis -
XX
XX PS      Disclosure; Fig 2; 54pp; English.
XX
XX CC      The invention relates to compound that competitively inhibits binding
XX CC      of competence signal peptide (CSP) to Streptococcus mutans histidine
XX CC      kinase (HK). Compounds of the invention are useful for treating or
XX CC      prophylaxis of caries or endocarditis. Vector comprising nucleic acid
XX CC      encoding S. mutans CSP is useful for producing recombinant host cell
XX CC      capable of expressing it. The recombinant host cell produced by the
XX CC      method is useful for expressing peptide in culture. S. mutans CSP DNA
XX CC      is useful for identifying nucleic acid molecules encoding CSP activated
XX CC      peptide. It is also useful as probes and in assays to identify
XX CC      antagonists or inhibitors of the peptides produced by the nucleic acid
XX CC      molecules. It is also useful for preparing vaccines for preventing or
XX CC      treating the above mentioned conditions. Antibodies against CSP activity
XX CC      are also useful for preventing caries. The antibodies are also useful
XX CC      for screening organisms or tissues containing CSP peptide or CSP-like
XX CC      peptide and for immunopurifying the peptides. The CSP nucleic acid
XX CC      molecules are useful in assays for genetic competence. The present
XX CC      sequence is S. mutans comC CSP precursor gene.
XX
XX SQ      Sequence 141 BP; 58 A; 19 C; 25 G; 39 T; 0 other;

Alignment Scores:
Pred. No.:      2,59e-28      Length:      141
Score:      227.00      Matches:      46
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:      24      Gaps:      0

US-09-833-017B-2 (1-46) x AAD32791 (1-141)
QY      1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db      1 ATGAAAAAACACTATCATTAATAAAATGACTTTAAAGAAATTAAGACTCATGAATTAGAG 60
QY      21 IleIleIleGlySerGlySerLeuSerThrPhePheAtgLeuPheAsnArgSerPhe 40
Db      61 ATTATCATTTGGCGGAAGCGGAAGCCCTATCAACATTTTCCGGCTGTTTAAACAGAAGTTT 120
QY      41 ThrGlnAlaLeuGlyLys 46
Db      121 ACACAGCTTTGGGAAA 138
RESULT 2
AAD32898
ID      AAD32898 standard; DNA; 141 BP.
XX
XX AC      AAD32898;
```

```
XX
XX DT      01-JUL-2002 (first entry)
XX
XX DE      Streptococcus mutans ComC gene.
XX
XX KW      Competence signal peptide; CSP; histidine kinase; HK; prophylaxis;
XX KW      therapy; caries; endocarditis; microbial biofilm; infection; gene;
XX KW      vaccine; antibacterial; ComC gene; ds.
XX
XX OS      Streptococcus mutans.
XX
XX FH      Key      Location/Qualifiers
XX FT      CDS      1..141
XX FT      /*tag= a
XX FT      /product= "S. mutans ComC protein"
XX
XX PN      CA2332733-A1.
XX
XX PD      10-OCT-2001.
XX
XX PF      20-FEB-2001; 2001CA-2332733.
XX
XX PR      10-APR-2000; 2000CA-2302861.
XX
XX PA      (LIYH/) LI Y H.
XX PA      (CVIT/) CVITKOVITCH D G.
XX PA      (LAUP/) LAU P C.
XX
XX PI      Li YH, Cvitkovitch DG, Lau PC;
XX
XX WPI: 2002-242173/30.
XX P-PSDB; AAE20594.
XX
XX PT      Novel compound that competitively inhibits binding of competence signal
XX PT      peptide to Streptococcus mutans histidine kinase, useful in treatment
XX PT      or prophylaxis of caries or endocarditis -
XX
XX PS      Disclosure; Fig 2; 82pp; English.
XX
XX CC      The invention relates to a compound that competitively inhibits the
XX CC      binding of competence signal peptide (CSP) to Streptococcus mutans
XX CC      histidine kinase (HK). Compounds of the invention are useful in
XX CC      medical treatment or prophylaxis of caries or endocarditis. They are
XX CC      useful for inhibiting or disrupting microbial biofilms involved in
XX CC      infections in man and animals and in biofouling of surfaces susceptible
XX CC      to microbial accumulation. They are also useful for treatment or
XX CC      prophylaxis of a disease, disorder or abnormal physical state caused
XX CC      by S. mutans. Pharmaceutical composition containing the compounds of
XX CC      the invention is useful for treating diseases caused by streptococcal
XX CC      infections. Polynucleotides encoding S. mutans CSP are useful as probes
XX CC      or in assays to identify antagonists or inhibitors of CSP peptides.
XX CC      They are also used for preparing vaccines. Polypeptides of the invention
XX CC      are useful for preparing antibodies, for vitro analysis of HK, CSP or
XX CC      RR activity or structure, and in assays for the identification and
XX CC      developments of compounds to inhibit and/or enhance polypeptide or
XX CC      peptide function directly. Antibodies of the invention are useful for
XX CC      providing protection against caries, to screen organisms or tissues
XX CC      containing CSP peptide or CSP-like peptides, for immuno-purification of
XX CC      CSP or CSP-like peptides from crude extracts, and to detect CSP or a
XX CC      similar peptide. The present sequence is S. mutans ComC gene encoding a
XX CC      CSP precursor protein.
XX
XX SQ      Sequence 141 BP; 58 A; 19 C; 25 G; 39 T; 0 other;

Alignment Scores:
Pred. No.:      2,59e-28      Length:      141
Score:      227.00      Matches:      46
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:      24      Gaps:      0

US-09-833-017B-2 (1-46) x AAD32898 (1-141)
```

QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
 Db 1 ATGAAAAAACACTATCATTAATAATGACITTAAGAATTAAGACTGATGAATTAGAG 60
 QY 21 IleIleIleGlySerGlySerLeuSerThrPheArgLeuPheAsnArgSerPhe 40
 Db 61 ATTATCATTTGGCGGAAGCGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTTT 120
 QY 41 ThrGlnAlaLeuGlyLys 46
 Db 121 ACACAAGCTTTGGGAAA 138
 RESULT 3
 ID AAD32800 standard; DNA; 2557 BP.
 XX
 AC AAD32800;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Streptococcus mutans comCDE gene local region.
 XX
 KW Competence signal peptide; CSP; histidine kinase; HK; therapy; carries;
 KW endocarditis; immunopurification; antibacterial; antiinflammatory;
 KW genetic competence assay; vaccine; comCDE gene; gene; ds.
 XX
 OS Streptococcus mutans.
 XX
 FH Location/Qualifiers
 FT complement (25..168)
 FT CDS
 FT /*tag= a
 FT /product= "Protein #1 encoded by S. mutans comCDE gene
 FT local region"
 FT complement (452..499)
 FT CDS
 FT /*tag= b
 FT /product= "Peptide #1 encoded by S. mutans comCDE gene
 FT local region"
 FT /note= "CDS does not include start and stop codon"
 FT partial
 FT 648..758
 FT CDS
 FT /*tag= c
 FT /product= "Peptide #2 encoded by S. mutans comCDE gene
 FT local region"
 FT 953..1081
 FT CDS
 FT /*tag= d
 FT /product= "Protein #2 encoded by S. mutans comCDE gene
 FT local region"
 FT complement (1366..1449)
 FT CDS
 FT /*tag= e
 FT /product= "Peptide #3 encoded by S. mutans comCDE gene
 FT local region"
 FT 1855..1959
 FT CDS
 FT /*tag= f
 FT /product= "Peptide #4 encoded by S. mutans comCDE gene
 FT local region"
 FT complement (1896..2072)
 FT CDS
 FT /*tag= g
 FT /product= "Protein #3 encoded by S. mutans comCDE gene
 FT local region"
 FT 2182..2424
 FT CDS
 FT /*tag= h
 FT /product= "Protein #4 encoded by S. mutans comCDE gene
 FT local region"
 FT 2384..2488
 FT CDS
 FT /*tag= i
 FT /product= "Peptide #5 encoded by S. mutans comCDE gene
 FT local region"
 FT
 XX CA2302861-A1.
 XX
 PN 10-OCT-2001.
 XX
 PD

PF 10-APR-2000; 2000CA-2302861.
 XX
 PR 10-APR-2000; 2000CA-2302861.
 XX
 PA (LAUP/) LAU P C Y.
 PA (CVIT/) CVITKOVIATCH D G.
 PA (LIYH/) LI Y H.
 XX
 PI Lau PCY, Cvitkovitch DG, Li YH;
 XX
 WPI: 2002-242151/30.
 DR P-PSDB; AAE20607, AAE20608, AAE20609, AAE20610, AAE20611, AAE20612,
 DR AAE20613, AAE20614, AAE20615.
 XX
 PT Novel compound that inhibits binding of competence signal peptide of
 PT Streptococcus mutans to S. mutans histidine kinase, useful for treating
 PT or preventing carries or endocarditis
 XX
 PS Disclosure; Fig 9; 54pp; English.
 XX
 CC The invention relates to compound that competitively inhibits binding
 CC of competence signal peptide (CSP) to Streptococcus mutans histidine
 CC kinase (HK). Compounds of the invention are useful for treating or
 CC prophylaxis of carries or endocarditis. Vector comprising nucleic acid
 CC encoding S. mutans CSP is useful for producing recombinant host cell
 CC capable of expressing it. The recombinant host cell produced by the
 CC method is useful for expressing peptide in culture. S. mutans CSP DNA
 CC is useful for identifying nucleic acid molecules encoding CSP activated
 CC peptide. It is also useful as probes and in assays to identify
 CC antagonists or inhibitors of the peptides produced by the nucleic acid
 CC molecules. It is also useful for preparing vaccines for preventing or
 CC treating the above mentioned conditions. Antibodies against CSP activity
 CC are also useful for preventing carries. The antibodies are also useful
 CC for screening organisms or tissues containing CSP peptide or CSP-like
 CC peptide and for immunopurifying the peptides. The CSP nucleic acid
 CC molecules are useful in assays for genetic competence. The present
 CC sequence is S. mutans comCDE gene local region.
 XX
 SQ Sequence 2557 BP; 891 A; 415 C; 426 G; 825 T; 0 other;
 XX
 Alignment Scores:
 Pred. NO.: 1.27e-26 Length: 2557
 Score: 227.00 Matches: 46
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-833-017B-2 (1-46) x AAD32800 (1-2557)
 QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
 Db 101 ATGAAAAAACACTATCATTAATAATGACITTAAGAATTAAGACTGATGAATTAGAG 160
 QY 21 IleIleIleGlySerGlySerLeuSerThrPheArgLeuPheAsnArgSerPhe 40
 Db 161 ATTATCATTTGGCGGAAGCGAAGCCTATCAACATTTTCCGGCTGTTTAACAGAAGTTT 220
 QY 41 ThrGlnAlaLeuGlyLys 46
 Db 221 ACACAAGCTTTGGGAAA 238
 RESULT 4
 ID AAD32893 standard; DNA; 2557 BP.
 XX
 AC AAD32893;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Streptococcus mutans comCDE gene local region.
 XX
 KW Competence signal peptide; CSP; histidine kinase; HK; prophylaxis;

therapy; caries; endocarditis; microbial biofilm; infection; gene; vaccine; antibacterial; comCDE gene local region; ds.

Streptococcus mutans.

Key Location/Qualifiers
CDS complement (25..168)
/*tag= a
/product= "Protein #1 encoded by S. mutans ComCDE gene local region"
101..241
/*tag= b
/product= "S. mutans ComC protein"
complement (383..1708)
/*tag= c
/product= "S. mutans ComD protein"
complement (452..499)
/*tag= d
/product= "Peptide #1 encoded by S. mutans ComCDE gene local region"
/note= "CDS does not include start and stop codon"
partial
648..758
/*tag= e
/product= "Peptide #2 encoded by S. mutans ComCDE gene local region"
953..1081
/*tag= f
/product= "Protein #2 encoded by S. mutans ComCDE gene local region"
complement (1366..1449)
/*tag= g
/product= "Peptide #3 encoded by S. mutans ComCDE gene local region"
complement (1705..2457)
/*tag= h
/product= "S. mutans ComE protein"
1855..1959
/*tag= i
/product= "Peptide #4 encoded by S. mutans ComCDE gene local region"
complement (1896..2072)
/*tag= j
/product= "Protein #3 encoded by S. mutans ComCDE gene local region"
2182..2424
/*tag= k
/product= "Protein #4 encoded by S. mutans ComCDE gene local region"
2384..2488
/*tag= l
/product= "Peptide #5 encoded by S. mutans ComCDE gene local region"

CA2332733-A1.

10-OCT-2001.

20-FEB-2001; 2001CA-2332733.

10-APR-2000; 2000CA-2302861.

(LIYH/) LI Y H.
(CVIT/) CVITKOVITCH D G.
(LAUF/) LAU P C.

Li YH, Cvitkovitch DG, Lau PC;

WPI; 2002-242173/30.
P-PSDB; AAE20617, AAE20618, AAE20619, AAE20620, AAE20621, AAE20622, AAE20623, AAE20624, AAE20625, AAE20594, AAE20585, AAE20586.

Novel compound that competitively inhibits binding of competence signal

peptide to Streptococcus mutans histidine kinase, useful in treatment or prophylaxis of caries or endocarditis -
PS Disclosure; Fig 9; 82pp; English.

XX The invention relates to a compound that competitively inhibits the binding of competence signal peptide (CSP) to Streptococcus mutans histidine kinase (HK). Compounds of the invention are useful in medical treatment or prophylaxis of caries or endocarditis. They are useful for inhibiting or disrupting microbial biofilms involved in infections in man and animals and in biofouling of surfaces susceptible to microbial accumulation. They are also useful for treatment or prophylaxis of a disease, disorder or abnormal physical state caused by S. mutans. Pharmaceutical composition containing the compounds of the invention is useful for treating diseases caused by streptococcal infections. Polynucleotides encoding S. mutans CSP are useful as probes or in assays to identify antagonists or inhibitors of CSP peptides. CC They are also used for preparing vaccines. Polypeptides of the invention CC are useful for preparing antibodies, for *in vitro* analysis of HK, CSP or CC RR activity or structure, and in assays for the identification and CC developments of compounds to inhibit and/or enhance polypeptide or CC peptide function directly. Antibodies of the invention are useful for CC providing protection against caries, to screen organisms or tissues CC containing CSP peptide or CSP-like peptides, for immuno-purification of CC CSP or CSP-like peptides from crude extracts, and to detect CSP or a CC similar peptide. The present sequence is S. mutans comCDE gene local CC region.

XX Sequence 2557 BP; 891 A; 415 C; 426 G; 825 T; 0 other;

Alignment Scores:

Pred. No.: 1.27e-26 Length: 2557
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-833-017b-2 (1-46) x AAD32893 (1-2557)

Qy 1 MetLysLysThrLeuSerLeuLysAsnAppheLysGluIleLysThrAspGluLeuGlu 20
Db 101 ATGAAAAAACACTATCATTTAAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAGAG 160
Qy 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 161 ATTATCATTCGCGAAGCGGAGCCATCATCAACATTTTCGGCTGTTTACAGAAAGTTT 220

Qy 41 ThrGlnAlaLeuGlyLys 46

Db 221 ACACAGCTTTGGGAAAA 238

RESULT 5

AAD32790

ID AAD32790 standard; DNA; 63 BP.

XX AAD32790;

DT 01-JUL-2002 (first entry)

DE Streptococcus mutans competence signal peptide (CSP) DNA.

XX Competence signal peptide; CSP; histidine kinase; HK; therapy; caries;

XX endocarditis; immunopurification; antibacterial; antiinflammatory;

XX genetic competence assay; vaccine; gene; ds.

XX Streptococcus mutans.

XX Key Location/Qualifiers

XX CDS 1..63

XX /*tag= a

XX /product= "S. mutans competence signal peptide (CSP)"

XX /note= "CDS does not include start and stop codon"

FT XX /partial
 PN CA2302861-A1.
 XX
 PD 10-OCT-2001.
 XX
 PF 10-APR-2000; 2000CA-2302861.
 XX
 PR 10-APR-2000; 2000CA-2302861.
 XX
 PA (LAUP/) LAU P C Y.
 PA (CVIT/) CVITKOVITCH D G.
 PA (LIYH/) LI Y H.
 XX
 PI Lau PCY, Cvitkovitch DG, Li YH;
 XX
 DR WPI: 2002-242151/30.
 DR P-PSDB: AAE20493.
 XX
 XX Novel compound that inhibits binding of competence signal peptide of Streptococcus mutans to S. mutans histidine kinase, useful for treating or preventing caries or endocarditis
 PT
 PT
 PS Claim 8; Fig 2; 54pp; English.
 XX
 CC The invention relates to compound that competitively inhibits binding of competence signal peptide (CSP) to Streptococcus mutans histidine kinase (HK). Compounds of the invention are useful for treating or prophylaxis of caries or endocarditis. Vector comprising nucleic acid encoding S. mutans CSP is useful for producing recombinant host cell capable of expressing it. The recombinant host cell produced by the method is useful for expressing peptide in culture. S. mutans CSP DNA is useful for identifying nucleic acid molecules encoding CSP activated peptide. It is also useful as probes and in assays to identify antagonists or inhibitors of the peptides produced by the nucleic acid molecules. It is also useful for preparing vaccines for preventing or treating the above mentioned conditions. Antibodies against CSP activity are also useful for preventing caries. The antibodies are also useful for screening organisms or tissues containing CSP peptide or CSP-like peptide and for immunopurifying the peptides. The CSP nucleic acid molecules are useful in assays for genetic competence. The present sequence is S. mutans CSP DNA.
 XX
 SQ Sequence 63 BP; 20 A; 12 C; 13 G; 18 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,01e-08 Length: 63
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.81% Indels: 0
 DB: 24 Gaps: 0
 US-09-833-017B-2 (1-46) x AAD32790 (1-63)
 QY 26 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 45
 Db 1 AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAAACAGAAGTTTACACAAGCTTTGGGA 60
 QY 46 Lys 46
 Db 61 AAA 63
 RESULT 6
 AAD32884
 ID AAD32884 standard; DNA; 63 BP.
 XX
 AC AAD32884;
 XX
 DF 01-JUL-2002 (first entry)
 XX Streptococcus mutans competence signal peptide (CSP) DNA.
 DE
 XX

KW Competence signal peptide: CSP; histidine kinase; HK; prophylaxis; therapy; caries; endocarditis; microbial biofilm; infection; gene; vaccine; antibacterial; ds.
 KW Streptococcus mutans.
 OS
 XX
 XX Location/Qualifiers
 FH 1..63
 FT /*tag= a
 FT /product= "S. mutans competence signal peptide"
 FT /note="CDS does not include start and stop codon"
 FT /partial
 XX
 XX CA2332733-A1.
 XX
 XX 10-OCT-2001.
 XX
 XX 20-FEB-2001; 2001CA-2332733.
 XX
 XX 10-APR-2000; 2000CA-2302861.
 XX
 XX (LIYH/) LI Y H.
 PA (CVIT/) CVITKOVITCH D G.
 PA (LAUP/) LAU P C.
 XX
 PI Li YH, Cvitkovitch DG, Lau PC;
 XX
 XX WPI: 2002-242173/30.
 DR P-PSDB: AAE20584.
 XX
 XX Novel compound that competitively inhibits binding of competence signal peptide to Streptococcus mutans histidine kinase, useful in treatment or prophylaxis of caries or endocarditis
 PT
 PT
 PS Claim 8; Fig 2; 82pp; English.
 XX
 CC The invention relates to a compound that competitively inhibits the binding of competence signal peptide (CSP) to Streptococcus mutans histidine kinase (HK). Compounds of the invention are useful in medical treatment or prophylaxis of caries or endocarditis. They are useful for inhibiting or disrupting microbial biofilms involved in infections in man and animals and in biofouling of surfaces susceptible to microbial accumulation. They are also useful for treatment or prophylaxis of a disease, disorder or abnormal physical state caused by S. mutans. Pharmaceutical composition containing the compounds of the invention is useful for treating diseases caused by streptococcal infections. Polynucleotides encoding S. mutans CSP are useful as probes or in assays to identify antagonists or inhibitors of CSP peptides. They are also used for preparing vaccines. Polypeptides of the invention are useful for preparing antibodies, for vitro analysis of HK, CSP or RR activity or structure, and in assays for the identification and developments of compounds to inhibit and/or enhance polypeptide or peptide function directly. Antibodies of the invention are useful for providing protection against caries, to screen organisms or tissues containing CSP peptide or CSP-like peptides, for immuno-purification of CSP or CSP-like peptides from crude extracts, and to detect CSP or a similar peptide. The present sequence is S. mutans CSP DNA.
 XX
 SQ Sequence 63 BP; 20 A; 12 C; 13 G; 18 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,01e-08 Length: 63
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 45.81% Indels: 0
 DB: 24 Gaps: 0
 US-09-833-017B-2 (1-46) x AAD32884 (1-63)
 QY 26 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 45
 Db 1 AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAAACAGAAGTTTACACAAGCTTTGGGA 60

Qy	46 Lys 46 	DB:	19	Gaps:	2
Db	61 AAA 63	US-09-833-017B-2 (1-46) x AAV52149 (1-8411)			
RESULT 7					
AAV52149/C					
ID	AAV52149 standard; DNA; 8411 BP.				
XX					
AC	AAV52149;				
DT	23-OCT-1998 (first entry)				
XX					
DE	Streptococcus pneumoniae genome fragment SEQ ID NO:16.				
XX					
KW	Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;				
KW	computer readable medium; vaccine; pharmaceutical composition; ds.				
XX					
OS	Streptococcus pneumoniae.				
XX					
PN	WO9818931-A2.				
XX					
PD	07-MAY-1998.				
XX					
PF	30-OCT-1997; 97WO-US19588.				
XX					
PR	31-OCT-1996; 96US-0029960.				
XX					
PA	(HUMA-) HUMAN GENOME SCI INC.				
XX					
PI	Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;				
PI	Kunsch CA, Rosen CA;				
XX					
DR	WPI; 1998-272225/24.				
XX					
PT	Computer-readable medium with recorded Streptococcus pneumoniae				
PT	polynucleotide sequences - useful in diagnostic kits and assays, and				
PT	pharmaceutical compositions and vaccines for Streptococcus				
PT	pneumoniae				
XX					
PS	Claim 1; Page 239-244; 1409pp; English.				
XX					
CC	The present invention describes a computer readable medium which has				
CC	the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)				
CC	recorded on it, or a representative fragment or a sequence at least 95%				
CC	identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in				
CC	SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from				
CC	Streptococcus pneumoniae. The present invention also describes an				
CC	isolated nucleic acid molecule encoding a homologue of any of the				
CC	fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the				
CC	nucleic acid molecule is produced by a process comprising: (a) screening				
CC	a genomic DNA library using as a probe a target sequence defined by any				
CC	of the sequences in SEQ ID NO:1 to 391, identifying members of the				
CC	library which contain sequences that hybridise to the target sequence and				
CC	isolating the nucleic acid molecules from the members; or (b) isolating				
CC	mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid				
CC	molecules whose nucleotide sequence is homologous to amplification				
CC	primers derived from the fragment of the S. pneumoniae genome to prime				
CC	the amplification and isolating the amplified sequences. The computer				
CC	readable medium can be used in a computer-based system for identifying				
CC	fragments of the S. pneumoniae genome of commercial importance, or				
CC	expression modulating fragments of the S. pneumoniae genome. Products				
CC	from the present invention can be used in diagnosis kits and assays, and				
CC	pharmaceutical compositions and vaccines for S. pneumoniae.				
XX					
SQ	Sequence 8411 BP; 2741 A; 1386 C; 1639 G; 2645 T; 0 other;				
Alignment Scores:					
Pred. No.:	3.67	Length:	8411		
Score:	69.00	Matches:	19		
Percent Similarity:	50.88%	Conservative:	10		
Best Local Similarity:	33.33%	Mismatches:	16		
Query Match:	30.40%	Indels:	12		

CC the disclosure of the invention.

XX Sequence 681 BP; 69 A; 101 C; 283 G; 228 T; 0 other;

Alignment Scores:

Pred. No.: 5.38 Length: 681
Score: 59.00 Matches: 11
Percent Similarity: 60.00% Conservative: 4
Best Local Similarity: 44.00% Mismatches: 10
Query Match: 25.99% Indels: 0
DB: 24 Gaps: 0

US-09-833-017B-2 (1-45) x ABQ15020 (1-681)

QY 21 llellelleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 255 TTGTTGTTGGCGGCGGCGTGTAGTTTCGGGTTTTTTTCGGGCGTTTCGTCGTCGTC 314

QY 41 ThrGlnAlaLeuGly 45

Db 315 GCGCGGGTATGGA 329

RESULT 9

ABQ15021/c

ID ABQ15021 standard; DNA; 681 BP.

XX AC ABQ15021;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 1612.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP10074.

XX PR 01-SEP-2000; 2000DE-1043826.

XX PR 05-SEP-2000; 2000DE-1044543.

XX PA (EP1G-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful
XX PT for diagnosis and prognosis, comprises selective hybridization of
XX PT amplicons from chemically treated DNA

XX PS Claim 12: 56pp + sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one
XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX and the degree of hybridisation to both classes is determined from the
XX label on the amplicon. From the ratio of labels hybridised to the two
XX classes of oligomers, the degree of methylation is calculated. The method
XX is used: (1) for diagnosis and/or prognosis of side effects of
XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX of the central nervous, cardiovascular, gastrointestinal and respiratory

CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.

XX SQ Sequence 681 BP; 228 A; 283 C; 101 G; 69 T; 0 other;

Alignment Scores:

Pred. No.: 5.38 Length: 681
Score: 59.00 Matches: 11
Percent Similarity: 60.00% Conservative: 4
Best Local Similarity: 44.00% Mismatches: 10
Query Match: 25.99% Indels: 0
DB: 24 Gaps: 0

US-09-833-017B-2 (1-46) x ABQ15021 (1-681)

QY 21 llellelleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 427 TTGTTGTTGGCGGCGGCGTGTAGTTTCGGGTTTTTTTCGGGCGTTTCGTCGTCGTC 368

QY 41 ThrGlnAlaLeuGly 45

Db 367 GCGCGGGTATGGA 353

RESULT 10

AAH87647/c

ID AAH87647 standard; cDNA; 1127 BP.

XX AC AAH87647;

XX DT 25-SEP-2001 (first entry)

XX DE Peppermint plant oil gland expressed cDNA 3.

XX KW Peppermint; plant oil gland cell; terpenoid essential oil; resin;
XX KW genetic mapping; antisense suppression; recombinant expression; ss.

XX OS Mentha x piperita.

XX PN WO200153319-A1.

XX PD 26-JUL-2001.

XX PF 19-JAN-2001; 2001WO-US02567.

XX PR 20-JAN-2000; 2000US-0177264.

XX PA (CROT/) CROTEAU R B.

XX PA (LANG/) LANGE B M.

XX PA (WILD/) WILDUNG M R.

XX PI Croteau RB, Lange BM, Wildung MR;

XX DR WPI; 2001-488706/53.

XX New nucleic acid molecules corresponding to mRNA molecules expressed in
XX peppermint oil glands for enhancing expression of plant oil gland cell
XX proteins

XX Claim 1: Page 77; 251pp; English.

XX The invention relates to nucleic acid molecules (AAH87645-AAH88116) that
XX correspond to all or part of a mRNA molecule expressed in plant oil
XX gland cells, especially peppermint and plant oil glands that produce
XX terpenoid essential oils and resins. The nucleic acids are useful for
XX genetically mapping a plant genome for genes expressed in plant oil
XX gland cells and to suppress (for example by antisense suppression) or
XX enhance their expression (for example by genetically transforming a
XX plant cell with a replicable expression vector that expresses one or more

CC proteins naturally expressed in plant oil gland cells). The nucleic acids
CC are also useful for recombinant expression of plant oil gland proteins
CC required for terpenoid essential oil and/or resin production in bacterial
CC and/or yeast cells.

SQ Sequence 1127 BP; 308 A; 244 C; 247 G; 327 T; 1 other;

Alignment Scores: Pred. No.: 10.6 Length: 1127
Score: 59.00 Matches: 17
Percent Similarity: 56.10% Conservative: 6
Best Local Similarity: 41.46% Mismatches: 14
Query Match: 25.99% Indels: 4
DB: 22 Gaps: 1

US-09-833-017B-2 (1-46) x AAF87647 (1-1127)

Qy 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 334 CTTAAGAAACATTATCTTGCAGCGAA-----AAAGAGTGAGTCTTGAA 287

Qy 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 286 ATCCAGCCCTCCATCCACTTCATCAGCTAATTTGTTTAATTATTCTCAAGATCATAT 227

Qy 41 Thr 41

Db 226 ACC 224

RESULT 11

AAAF85076/c

ID AAF85076 standard; DNA; 1481 BP.

XX AAF85076;

09-JUL-2001 (first entry)

Nucleotide sequence of an isopentenyl monophosphatase kinase.

XX Isopentenyl monophosphate kinase; IPK; isoprenoid biosynthesis; pigment;
KW vitamin; essential oil; pathogen resistance; antibiotic; herbicide;
KW antimalarial; ss.

XX Mentha piperita.

XX Key Location/Qualifiers

FT CDS 3..1220

FT /*tag= a /product= "isopentenyl monophosphatase kinase"

XX WO200132907-A1.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30289.

XX 04-NOV-1999; 99US-0434774.

XX (UNIW) UNIV WASHINGTON STATE RES FOUND.

XX Croteau RB, Lange BM;

XX WPI; 2001-308747/32.

XX P-PSDB; AAB68285.

XX New nucleic acid encoding isopentenyl monophosphate kinase, useful e.g.
PT for making transgenic plants with increased synthesis of isoprenoids,
PT e.g. essential oils -

XX Claim 12; Page 42-44; 62pp; English.

XX The present sequence encodes an isopentenyl monophosphate kinase (IPK)
CC polypeptide. The enzyme is an isoprenoid biosynthesis stimulator. The

CC IPK polynucleotide is useful recombinant production of IPK, as a source
CC of probes, primers and antisense sequences and for increasing/reducing
CC expression levels of IPK in cells, particularly of essential oil
CC plants, so as to increase flow through the isoprenoid biosynthesis
CC pathway, resulting in increased production of e.g. pigments, vitamins
CC and essential oils, also to increase resistance to pests and pathogens
CC and to improve plant strength. Mutant forms of the IPK polynucleotide
CC can be used to express forms of IPK that are resistant to IPK-targeted
CC herbicides, and recombinant IPK can be used to screen for antibiotics,
CC herbicides and antimalarial agents directed against IPK.

XX SQ Sequence 1481 BP; 413 A; 320 C; 302 G; 446 T; 0 other;

Alignment Scores: Pred. No.: 15.3 Length: 1481
Score: 59.00 Matches: 17
Percent Similarity: 56.10% Conservative: 6
Best Local Similarity: 41.46% Mismatches: 14
Query Match: 25.99% Indels: 4
DB: 22 Gaps: 1

US-09-833-017B-2 (1-46) x AAF85076 (1-1481)

Qy 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 334 CTTAAGAAACATTATCTTGCAGCGAA-----AAAGAGTGAGTCTTGAA 287

Qy 21 IleIleIleGlyGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 286 ATCCAGCCCTCCATCCACTTCATCAGCTAATTTGTTTAATTATTCTCAAGATCATAT 227

Qy 41 Thr 41

Db 226 ACC 224

RESULT 12

ABK77550

ID ABK77550 standard; DNA; 416 BP.

XX ABK77550;

XX 13-AUG-2002 (first entry)

XX Bacillus clausii genomic sequence tag (GST) #393.

XX Differential gene expression; genomic sequenced tag; GST;

KW altered culture condition; environmental stress;

KW physiological provocation; ds.

XX Bacillus clausii.

XX WO200229113-A2.

XX 11-APR-2002.

XX 05-OCT-2001; 2001WO-US31437.

XX 06-OCT-2000; 2000US-0680598.

XX 27-MAR-2001; 2001US-279526P.

XX (NOVO) NOVOZYMES BIOTECH INC.

XX Berka R, Clausen IG;

XX WPI; 2002-416684/44.

XX Monitoring differential expression of several genes in first Bacillus

PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array

XX Claim 11; SEQ ID NO 4841; 200pp; English.

XX The invention describes a method of monitoring differential expression of
 CC genes in a first Bacillus cell relative to expression of the genes in
 CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
 CC isolated from Bacillus cells to a substrate containing array of Bacillus
 CC genomic sequenced tags (GST), examining the array, and determining
 CC relative gene expression by an observed hybridisation reporter signal of
 CC a spot in the array. The method is useful for measuring the expression of
 CC genes in a first Bacillus cell relative to expression of the same genes
 CC in one or more second Bacillus cells. The method is useful for monitoring
 CC global expression of several genes from a Bacillus cell, discovering new
 CC genes, identifying possible functions of unknown open reading frames and
 CC monitoring gene copy number variation and stability. Monitoring changes
 CC in expression of genes may be used to provide a representation of the way
 CC in which Bacillus cells adapt to changes in culture conditions,
 CC environmental stress or other physiological provocation. Extensive
 CC follow-up characterisation is unnecessary, when one spot on an array
 CC equals one gene or one open reading frame, since sequence information is
 CC available. This sequence represents a genomic sequence tag (GST) used in
 CC the method of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 416 BP; 115 A; 74 C; 118 G; 109 T; 0 other;

Alignment Scores:
 Pred. No.: 3.35 Length: 416
 Score: 58.50 Matches: 18
 Percent Similarity: 60.53% Conservative: 5
 Best Local Similarity: 47.37% Mismatches: 13
 Query Match: 25.77% Indels: 2
 DB: 24 Gaps: 1

US-09-833-017B-2 (1-46) x ABK77550 (1-416)

QY 3 LysThrLeuSerLeuLysAsnAspPhe-LysGluIleLysThrAspGluLeuGluIle1 22

DB 36 AGAATCATGAAGTCTAAAGGGGATTACGCGGACATTCGCGAAAGGAAGTCTTTGTCAT 95

QY 22 elledGlySerGlySer---LeuSerThrPhePheArgLeuPheAsnArg 38

DB 96 TATTGCGCGTGTGCTGCTCGGAAAGTACGTTTCTCCGCTGTTTAAACCG 147

RESULT 13

ID ABQ46476

AC ABQ46476 standard; DNA; 885 BP.

XX ABQ46476;

DT 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 33067.

DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

OS Homo sapiens.

PN WO200218632-A2.

PD 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP10074.

XX 01-SEP-2000; 2000DE-1043826.

PR 05-SEP-2000; 2000DE-1044543.

XX (EPIG-) EPIGENOMICS AG.

PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful
 CC for diagnosis and prognosis, comprises selective hybridization of
 CC amplicons from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.

XX Sequence 885 BP; 127 A; 99 C; 327 G; 332 T; 0 other;

Alignment Scores:
 Pred. No.: 9.23 Length: 885
 Score: 58.50 Matches: 15
 Percent Similarity: 41.67% Conservative: 10
 Best Local Similarity: 25.00% Mismatches: 16
 Query Match: 25.77% Indels: 19
 DB: 24 Gaps: 1

US-09-833-017B-2 (1-46) x ABQ46476 (1-885)

QY 5 LeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGluIle----- 21

DB 89 TTGGCGTTTCGTTTCGTTTCGCGAGGTTTCGCGCGGTTTCGCGCGAGGTTGG 148

QY 22 -----IleIleGlyGly 25

DB 149 AGCGAGTTTCGGACGCGAGAGATTGGCGGATTGGTGGATTGGTAGACGTTGGCGGA 208

QY 26 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 45

DB 209 CGGGGTGATTAGGTACGTGCTTCGCGATGCTGTTACGATCGTTTCGTCGCTTAGGT 268

RESULT 14

ABQ46477/c

ID ABQ46477 standard; DNA; 885 BP.

XX ABQ46477;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 33068.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.

OS Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.
 XX 01-SEP-2001; 2001WO-EP10074.
 XX 01-SEP-2000; 2000DE-1043826.
 PR 05-SEP-2000; 2000DE-1044543.
 XX (EPIC-) EPIGENOMICS AG.
 XX Olek A, Plepenbrock C, Berlin K, Guetig D;
 XX WPI; 2002-371829/40.
 XX Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA
 XX
 XX Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNPs); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 XX
 XX Sequence 885 BP; 332 A; 327 C; 99 G; 127 T; 0 other;
 Alignment Scores:
 Pred. No.: 9.23 Length: 885
 Score: 58.50 Matches: 15
 Percent Similarity: 41.67% Conservative: 10
 Best Local Similarity: 25.00% Mismatches: 16
 Query Match: 25.77% Indels: 19
 DB: 24 Gaps: 1
 US-09-833-017B-2 (1-46) x ABQ46477 (1-885)
 QY 5 LeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGluLeu----- 21
 DB 797 TTGGCGTTTCGTTTCGTTTCGCGAGGTTTCGCGCGGTTTCGCGAGGTTGG 738
 QY 22 -----leuileGlyGly 25
 DB 737 AGCGGAGTTTCGCGCGCAGAGATTGGCGGATTGTTGATTGAGCGTTGGCGGA 678
 QY 26 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 45
 DB 677 CGGGGTGATTAGGACGTTTCGTTTCGCGATGTTGTTAGATCGTTTCGTCGTTTAGT 618
 RESULT 15
 ABN70018/C
 ID ABN70018 standard; DNA; 1530 BP.
 XX
 XX ABN70018;
 XX
 XX 01-JUL-2002 (first entry)
 XX

DE Streptococcus polynucleotide SEQ ID NO 7949.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial; gene;
 KW antinflammatory; infection; vaccine; meningitis; gene therapy; ds.
 XX
 OS Streptococcus pyogenes.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 XX 29-OCT-2001; 2001WO-GB04789.
 PF
 XX 27-OCT-2000; 2000GB-0026333.
 PR
 XX 24-NOV-2000; 2000GB-0028727.
 PR
 XX 07-MAR-2001; 2001GB-0005640.
 XX
 XX (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 PA
 XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 PI
 XX WPI: 2002-352536/38.
 DR
 XX P-PSDB; ABP29387.
 DR
 XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 XX Claim 7; Page 3923-3924; 4525pp; English.
 XX
 XX The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN65044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 XX Sequence 1530 BP; 438 A; 313 C; 341 G; 438 T; 0 other;
 Alignment Scores:
 Pred. No.: 19.2 Length: 1530
 Score: 58.50 Matches: 16
 Percent Similarity: 52.27% Conservative: 7
 Best Local Similarity: 36.36% Mismatches: 20
 Query Match: 25.77% Indels: 1
 DB: 24 Gaps: 1
 US-09-833-017B-2 (1-46) x ABN70018 (1-1530)
 QY 3 LysThrLeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGluLeu 22
 DB 993 AAAGTCTTGACATCAAGATTTCACACCAAGTCTCTTCGAGCAACTCCACCATAGC 934
 QY 23 IleGlyGly---SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThr 41
 DB 933 ATTGGTAATACCAAGCAAGCAAGTAGCTTTCATATTTCTTTATCATCATGCTTACG 874
 QY 42 GlnAlaLeuGly 45
 |||

Wed Nov 13 10:42:11 2002

us-09-833-017b-2.rng

Page 11

Db 873 CATGAAACGGGT 862

Search completed: November 8, 2002, 22:10:56
Job time : 233.313 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 22:05:37 : Search time 48.0597 Seconds
(without alignments)
293.533 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 227

Sequence: 1 MKKTLKNDKFKIKTDELE.....GSLSTFFRLNRSFTQALGK 46

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09833017/runat_05112002_105349_4858/app_query.fasta.1.398
-DB=Issued_Patents_NA -FAST=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017.ecgn.1.1.17 @runat_05112002_105349_4858 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1:	/cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2:	/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3:	/cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4:	/cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5:	/cgn2_6/ptodata/1/ina/PCITUS_COMB.seq:*
6:	/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	69	30.4	8411	4	US-08-961-527-16
C 2	59	26.0	1481	4	US-09-434-774-1
C 3	55.5	24.4	1662	2	US-08-663-566A-12
4	55.5	24.4	1662	2	US-08-023-610-12
5	55.5	24.4	1662	2	US-08-288-065A-12
6	55.5	24.4	1662	2	US-08-362-240A-12
7	55.5	24.4	1662	4	US-08-804-372A-10
8	55.5	24.4	1662	5	PCT-US95-10245-12
9	55.5	24.4	3825	6	5310678-2
10	55.5	24.4	4177	2	US-08-484-575A-12
11	55.5	24.4	4177	3	US-08-477-459-12
12	55.5	24.4	4177	3	US-08-479-869-12

13	55.5	24.4	4177	3	US-08-486-414-12	Sequence 12, Appl
14	55.5	24.4	4177	5	PCT-US94-01826A-12	Sequence 12, Appl
15	55.5	24.4	4177	5	PCT-US94-02252A-12	Sequence 12, Appl
16	55	24.0	3033	4	US-09-134-001C-2341	Sequence 2341, Ap
17	54.5	24.0	5848	4	US-09-368-588-3	Sequence 3, Appl
18	54	23.8	738	4	US-09-221-017B-246	Sequence 246, App
19	54	23.8	738	1	US-08-409-995-3	Sequence 3, Appl
20	54	23.8	738	3	US-08-685-467-3	Sequence 3, Appl
21	54	23.8	7253	4	US-09-268-347-35	Sequence 35, Appl
22	54	23.8	7291	4	US-08-913-942-3	Sequence 3, Appl
23	54	23.8	29793	4	US-09-302-812-38	Sequence 38, Appl
24	54	23.8	29793	4	US-09-511-477-38	Sequence 38, Appl
25	54	23.8	29793	4	US-09-511-507-38	Sequence 38, Appl
26	53.5	23.6	1197	4	US-08-745-995A-10	Sequence 10, Appl
27	53.5	23.6	1197	4	US-08-745-995A-11	Sequence 11, Appl
28	53.5	23.6	1358	4	US-08-745-995A-7	Sequence 9, Appl
29	53.5	23.6	1358	4	US-08-745-995A-9	Sequence 24, Appl
30	53.5	23.6	1845	3	US-08-989-251-24	Sequence 24, Appl
31	53.5	23.6	1845	3	US-09-340-250-24	Sequence 24, Appl
32	53.5	23.6	1845	3	US-09-528-108-24	Sequence 36, Appl
33	53.5	23.6	2023	3	US-08-989-251-36	Sequence 36, Appl
34	53.5	23.6	2023	3	US-09-340-250-36	Sequence 36, Appl
35	53.5	23.6	2023	4	US-09-528-108-36	Sequence 1, Appl
36	53	23.3	1260	2	US-08-578-158-1	Sequence 13, Appl
37	53	23.3	10523	4	US-09-453-702B-13	Sequence 11, Appl
38	53	23.3	12720	1	US-08-403-866-11	Sequence 2, Appl
39	52	22.9	861	1	US-07-877-516-2	Sequence 1, Appl
40	52	22.9	1242	1	US-07-877-516-1	Sequence 1, Appl
41	52	22.9	1947	3	US-08-604-991-1	Sequence 1, Appl
42	52	22.9	1947	3	US-09-363-639-1	Sequence 1, Appl
43	52	22.9	8959	1	US-08-920-812-1	Sequence 1, Appl
44	52	22.9	8959	1	US-08-920-827-1	Sequence 1, Appl
45	52	22.9	8959	1	US-08-921-177-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-961-527-16/c

; Sequence 16, Application US/08961527

; Patent No. 6420135

; GENERAL INFORMATION:

; APPLICANT: Charles Kunsch

; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences

; NUMBER OF SEQUENCES: 391

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/961.527

; FILING DATE:

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Brookes, A. Anders

; REGISTRATION NUMBER: 36,373

; REFERENCE/DOCKET NUMBER: PB340P1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 16:

; SEQUENCE CHARACTERISTICS:

```

; LENGTH: 8411 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-16

Alignment Scores:
Pred. No.: 0.752 Length: 8411
Score: 69.00 Matches: 19
Percent Similarity: 50.88% Conservative: 10
Best Local Similarity: 33.33% Mismatches: 16
Query Match: 30.40% Indels: 12
DB: 4 Gaps: 2

US-09-833-017B-2 (1-46) x US-08-961-527-16 (1-8411)
QY 2 LysLysThrLeuSerLeuLysAsnAspPhe-----LysGlu 13
||||| :||| ||| |||
Db 3566 AAAAGACATTACAACTGATATGCTTCCTCTTTATCTAGGTACTAAAGAG 3507
QY 14 IleLysThrAspGluLeuLlelleleGlyGlySerGlySerLeuSer----- 30
||||| :||| ||| |||
Db 3506 AAAAACCITTCATATCAATGTTAGTTCCCTCGAAATTCACAAGTATCATTTGAATCT 3447
QY 31 ---ThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 46
||||| :||| ||| |||
Db 3446 CCTACCTTCTATCGTTTAGATACCAAAACTTTCCACCGAGGCAATTCAAAAA 3396

RESULT 2
US-09-434-774-1/C
; Sequence 1, Application US/09434774A
; Patent No. 6235514
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Lange, Bernd M
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING ISOPENENTYL
; TITLE OF INVENTION: MONOPHOSPHATE KINASE, AND METHODS OF USE
; FILE REFERENCE: wsur14448
; CURRENT APPLICATION NUMBER: US/09/434,774A
; CURRENT FILING DATE: 1999-11-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Mentha piperita
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(1217)
US-09-434-774-1

Alignment Scores:
Pred. No.: 3.25 Length: 1481
Score: 59.00 Matches: 17
Percent Similarity: 56.10% Conservative: 6
Best Local Similarity: 41.46% Mismatches: 14
Query Match: 25.99% Indels: 4
DB: 4 Gaps: 1

US-09-833-017B-2 (1-46) x US-09-434-774-1 (1-1481)
QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluLeuLysThrAspGluLeuGlu 20
||||| :||| ||| |||
Db 334 CTTTAGAANAACATTAATCTTCGACGGCAA-----AAAAGAGTGAGTCTTGAA 287
QY 21 llellelleGlyGlySerGlySerThrPhePheArgLeuPheAsnArgSerPhe 40
|||| |||| ||| :||| ||| ||| ||| ||| ||| ||| ||| |||
Db 286 ATCCCGACGCTCCCTATCCACTTTCATGACGTAAATGTGTTAAATTTATTCAGATCATAT 227
QY 41 Thr 41
|||
Db 226 ACC 224

```


STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,240A
FILING DATE: Dec-22-94
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)278-0400
TELEFAX: (212)391-0526
TELEX: 422523
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1662
US-08-362-240A-12

Alignment Scores:
Pred. No.: 13.8 Length: 1662
Score: 55.50 Matches: 16
Percent Similarity: 47.92% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 20
Query Match: 24.45% Indels: 5
Gaps: 1

US-09-833-017B-2 (1-46) x US-08-362-240A-12 (1-1662)

QY 4 ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18
||||| ||| :||| ||| ||| :|||
Db 1297 ACTTTAAGCTCAGTGGGGAATTCGATTAATCAGAGAATATCTCAATACAAGAT 1356
QY 19 LeuGluIleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArg 38
:||||| ||| :||| :||| :||| :|||
Db 1357 TCTCAAGTAATAATAACAGGCAATCTTGATATCTCAACTGAGCTTGGGAATGTCAACAAC 1416
QY 39 SerPheThrGlnAlaLeuGlyLys 46
||| :||| :||| :||| :|||
Db 1417 TCGATCAGTAATGCCTTGAATAAG 1440

RESULT 7

US-08-804-372A-10
Sequence 10, Application US/08804372A
Patent No. 6183753
GENERAL INFORMATION:
APPLICANT: Cochran, Mark D.
APPLICANT: Wild, Martha A.
APPLICANT: Winslow, Barbara J.
TITLE OF INVENTION: Recombinant Chimeric Viruses and Uses
TITLE OF INVENTION: Thereof
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,372A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 2552/391115E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1662 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1662
US-08-804-372A-10

Alignment Scores:
Pred. No.: 13.8 Length: 1662
Score: 55.50 Matches: 16
Percent Similarity: 47.92% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 20
Query Match: 24.45% Indels: 5
Gaps: 1

US-09-833-017B-2 (1-46) x US-08-804-372A-10 (1-1662)

QY 4 ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18
||||| ||| :||| ||| ||| :|||
Db 1297 ACTTTAAGCTCAGTGGGGAATTCGATTAATCAGAGAATATCTCAATACAAGAT 1356
QY 19 LeuGluIleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArg 38
:||||| ||| :||| :||| :||| :|||
Db 1357 TCTCAAGTAATAATAACAGGCAATCTTGATATCTCAACTGAGCTTGGGAATGTCAACAAC 1416
QY 39 SerPheThrGlnAlaLeuGlyLys 46
||| :||| :||| :||| :|||
Db 1417 TCGATCAGTAATGCCTTGAATAAG 1440

RESULT 8

PCT-US95-10245-12
Sequence 12, Application PC/TUS9510245
GENERAL INFORMATION:
APPLICANT: SYNTRO CORPORATION
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof
NUMBER OF SEQUENCES: 60
CORRESPONDENCE ADDRESS:
ADDRESSEE: John P. White
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10245

```

; FILING DATE: 09-AUG-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1662 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1662
; PCT-US95-10245-12

Alignment Scores:
Pred. No.: 13.8 Length: 1662
Score: 55.50 Matches: 16
Percent Similarity: 47.92% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 20
Query Match: 24.45% Indels: 5
DB: 5 Gaps: 1

US-09-833-017B-2 (1-46) x PCT-US95-10245-12 (1-1662)

QY 4 ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18
||||| ||| :||| ||| |||
Db 1297 ACVTTAAGCTCAGTGGGGAATTCGATTACTATCAGAAGATAATCTCAATACAAGAT 1356

QY 19 LeuGluIleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArg 38
:::||||| |||::: |||
Db 1357 TCTCAAGTAATAATACAGGCAATCTTGATATCTCACTGAGTTGGGAATGTCAACAAC 1416

QY 39 SerPheThrGlnAlaLeuGlyLys 46
||| ::| ||||| |||
Db 1417 TCGATCAGTAATGCCCTTGAAATAG 1440

RESULT 9
5310678-2
; Patent No. 5310678
; APPLICANT: Bingham, Richard W.; Chambers, Philip; Emmerson, Peter
; T.: Miller, Neil S.
; TITLE OF INVENTION: NEWCASTLE DISEASE VIRUS GENE CLONES
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/438,945
; FILING DATE: 17-NOV-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 885,765
; FILING DATE: 15-JUL-1986
; SEQ ID NO: 2:
; LENGTH: 3825
5310678-2

Alignment Scores:
Pred. No.: 40.2 Length: 3825
Score: 55.50 Matches: 16
Percent Similarity: 47.92% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 20
Query Match: 24.45% Indels: 5
DB: 6 Gaps: 1

US-09-833-017B-2 (1-46) x 5310678-2 (1-3825)

QY 4 ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18

```

Db	3451	TCCTCAAGTAATAACAGGCAATCTTGATATCTCAACTGACGCTTGGGAATGTCAACAAC	3510
QY	39	SerPheThrGlnAlaLeuGlyLys 46	
Db	3511	TCGATCAGTAATGCCCTTGAATAAG	3534
RESULT 11			
US-08-477-459-12			
Sequence 12, Application US/08477459			
Patent No. 6001369			
GENERAL INFORMATION:			
APPLICANT: Mark D. Cochran			
TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses			
TITLE OF INVENTION: Thereof			
NUMBER OF SEQUENCES: 20			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: John P. White			
STREET: 1185 Avenue of the Americas			
CITY: New York			
STATE: New York			
COUNTRY: USA			
ZIP: 10036			
COMPUTER READABLE FORM:			
MEDIUM TYPE: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/477,459			
FILING DATE: 07-JUN-1995			
CLASSIFICATION: 424			
ATTORNEY/AGENT INFORMATION:			
NAME: White Esq, John P			
TELEPHONE: (212) 278-0400			
TELEFAX: (212) 391-0525			
INFORMATION FOR SEQ ID NO: 12:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 4177 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
HYPOTHETICAL: NO			
ANTI-SENSE: NO			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 115..1860			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 2095..3756			
US-08-477-459-12			
Alignment Scores:			
Pred. No.:	45	Length:	4177
Score:	55.50	Matches:	16
Percent Similarity:	47.92%	Conservative:	7
Best Local Similarity:	33.33%	Mismatches:	20
Query Match:	24.45%	Indels:	5
DB:	3	Gaps:	1
US-09-833-017B-2 (1-46) x US-08-477-459-12 (1-4177)			
QY	4	ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18	
Db	3391	ACTTTAAGCCTCAGTGGGAATTCGATTAATTATCAGAAGATATCTCAATACAAGAT 3450	
QY	19	LeuGluIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArg 38	
Db	3451	TCCTCAAGTAATAACAGGCAATCTTGATATCTCAACTGACGCTTGGGAATGTCAACAAC 3510	
QY	39	SerPheThrGlnAlaLeuGlyLys 46	

Db	3511	TCGATCAGTAATGCCCTTGAATAAG	3534
RESULT 12			
US-08-479-869-12			
Sequence 12, Application US/08479869			
Patent No. 6123949			
GENERAL INFORMATION:			
APPLICANT: Cochran Ph.D, Mark D			
TITLE OF INVENTION: Recombinant Fowlpox Virus S-FPV-043 and			
TITLE OF INVENTION: Uses Thereof			
NUMBER OF SEQUENCES: 20			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: John P. White			
STREET: 30 Rockefeller Plaza			
CITY: New York			
STATE: New York			
COUNTRY: USA			
ZIP: 10112			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.25			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/479,869			
FILING DATE:			
CLASSIFICATION: 424			
PRIOR APPLICATION DATA:			
APPLICATION NUMBER: 08/024,156			
FILING DATE: 26-FEB-1993			
ATTORNEY/AGENT INFORMATION:			
NAME: White Esq, John P			
TELEPHONE: (212)977-9550			
TELEFAX: (212)664-0525			
TELEX: 422523			
INFORMATION FOR SEQ ID NO: 12:			
SEQUENCE CHARACTERISTICS:			
LENGTH: 4177 base pairs			
TYPE: nucleic acid			
STRANDEDNESS: double			
TOPOLOGY: linear			
MOLECULE TYPE: DNA (genomic)			
HYPOTHETICAL: NO			
ANTI-SENSE: NO			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 115..1860			
FEATURE:			
NAME/KEY: CDS			
LOCATION: 2095..3756			
US-08-479-869-12			
Alignment Scores:			
Pred. No.:	45	Length:	4177
Score:	55.50	Matches:	16
Percent Similarity:	47.92%	Conservative:	7
Best Local Similarity:	33.33%	Mismatches:	20
Query Match:	24.45%	Indels:	5
DB:	3	Gaps:	1
US-09-833-017B-2 (1-46) x US-08-479-869-12 (1-4177)			
QY	4	ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu 18	
Db	3391	ACTTTAAGCCTCAGTGGGAATTCGATTAATTATCAGAAGATATCTCAATACAAGAT 3450	
QY	19	LeuGluIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArg 38	
Db	3451	TCCTCAAGTAATAACAGGCAATCTTGATATCTCAACTGACGCTTGGGAATGTCAACAAC 3510	
QY	39	SerPheThrGlnAlaLeuGlyLys 46	

; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4177 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..1860
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2095..3756
; PCT-US94-02252A-12

Alignment Scores:
Pred. No.: 45 Length: 4177
Score: 55.50 Matches: 16
Percent Similarity: 47.92% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 20
Query Match: 24.45% Indels: 5
DB: 5 Gaps: 1

US-09-833-017B-2 (1-46) x PCT-US94-02252A-12 (1-4177)

Qy	4	ThrLeuSerLeuLysAsnAspPhe-----LysGluIleLysThrAspGlu	18
Db	3391	ACTTTAAGGCTCAGTGGGAATTCGATGTAACCTATCAGAAGAATATCTCAATACAAGAT	3450
Qy	19	LeuGluIleIleIleGlySerGlySerLeuSerThrPheArgLeuPheAsnArg	38
Db	3451	TCTCAGTAAATTAACAGGCAATCTTGATATCTCACTGAGCTGGGAATGTCACAAAC	3510
Qy	39	SerPheThrGlnAlaLeuGlyLys	46
Db	3511	TCGATCAGTAATGCCTTGAATAAG	3534

Search completed: November 8, 2002, 23:41:15
Job time : 53.0597 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 22:57:42 ; Search time 45.3134 seconds
(without alignments)
360.161 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 227
Sequence: 1 MKKYLKLNDFKEIKTDELE.....GSLSTFFRLFNRSFTQALGK 46

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frames_p2n.model -DEV=xlh
-Q=/cgn2_1/USPro_spool/US09833017/runat_05112002_105351_4940/app_query.fasta_1.398
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomatch62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09833017@cgn.1.1.18 @runat_05112002_105351_4940
-NCPD=6 -ICPD=3 -NO_XLPXY -NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227	100.0	141	10	US-09-833-017-1 Sequence 1, Appli
2	227	100.0	2557	10	US-09-833-017-21 Sequence 21, Appl
3	58.5	25.8	416	10	US-09-974-300-4841 Sequence 4841, Ap
4	57	25.1	942	10	US-09-974-300-2095 Sequence 2095, Ap

c	5	57	25.1	48667	10	US-09-822-268A-3	Sequence 3, Appli
	6	56.5	24.9	843	10	US-09-974-300-3171	Sequence 3171, Ap
	7	56.5	24.9	13884	10	US-09-070-927A-341	Sequence 341, App
	8	56	24.7	2820	10	US-09-815-242-6700	Sequence 6700, Ap
	9	56	24.7	5480	10	US-09-070-927A-132	Sequence 132, App
c	10	56	24.7	465237	10	US-09-933-267A-1	Sequence 1, Appli
c	11	55.5	24.4	3570	10	US-09-881-457A-1	Sequence 1, Appli
c	12	55	24.2	3218	10	US-09-788-657-5	Sequence 5, Appli
c	13	54.5	24.0	11871	10	US-09-070-927A-171	Sequence 171, App
	14	54	23.8	2376	10	US-09-917-800A-1496	Sequence 1496, Ap
	15	54	23.8	25793	10	US-09-973-451-38	Sequence 38, Appli
c	16	53.5	23.6	1845	10	US-09-921-398-24	Sequence 24, Appli
c	17	53.5	23.6	2023	10	US-09-921-398-36	Sequence 36, Appli
c	18	53	23.3	438	10	US-09-924-035A-605	Sequence 605, App
c	19	53	23.3	971	10	US-09-770-445-285	Sequence 285, App
	20	52.5	23.1	24601	10	US-09-070-927A-223	Sequence 223, App
	21	52	22.9	269	10	US-09-294-093B-1839	Sequence 1839, Ap
	22	52	22.9	275	10	US-09-878-574-15440	Sequence 15440, A
c	23	52	22.9	10476	10	US-09-964-824A-98	Sequence 98, Appli
c	24	52	22.9	10476	10	US-09-964-824A-552	Sequence 552, App
	25	51.5	22.7	395	10	US-09-960-352-12338	Sequence 12338, A
	26	51.5	22.7	421	10	US-09-880-107-1416	Sequence 1416, Ap
	27	51.5	22.7	774	10	US-09-741-669-184	Sequence 184, App
	28	51.5	22.7	921	10	US-09-912-020-115	Sequence 115, App
	29	51.5	22.7	1827	10	US-09-887-576-817	Sequence 817, App
	30	51.5	22.7	2979	9	US-09-938-842A-48	Sequence 48, Appli
c	31	51.5	22.7	7237	10	US-09-070-927A-239	Sequence 239, App
	32	51.5	22.7	16870	10	US-09-070-927A-98	Sequence 98, Appli
	33	51.5	22.7	147309	10	US-09-742-312-3	Sequence 3, Appli
	34	51	22.5	263	10	US-09-878-574-15097	Sequence 15097, A
c	35	51	22.5	492	10	US-09-783-590-1335	Sequence 1335, Ap
	36	51	22.5	663	10	US-09-815-242-8789	Sequence 8789, Ap
c	37	51	22.5	1726	9	US-09-938-842A-4596	Sequence 4596, Ap
	38	51	22.5	2979	10	US-09-815-242-4748	Sequence 4748, Ap
c	39	51	22.5	12822	10	US-09-764-847-1579	Sequence 1579, Ap
	40	51	22.5	32193	10	US-09-764-877-2623	Sequence 2623, Ap
	41	50.5	22.2	333	10	US-09-815-242-4774	Sequence 4774, Ap
	42	50.5	22.2	360	10	US-09-815-242-8884	Sequence 8884, Ap
	43	50.5	22.2	360	10	US-09-815-242-9051	Sequence 9051, Ap
	44	50.5	22.2	408	10	US-09-878-574-3898	Sequence 3898, Ap
c	45	50.5	22.2	425	10	US-09-815-242-1718	Sequence 1718, Ap

ALIGNMENTS

RESULT 1
US-09-833-017-1
; Sequence 1, Application US/09833017
; Patent No. US20020081302A1
; GENERAL INFORMATION:
; APPLICANT: CVITKOVITCH, Dennis
; TITLE OF INVENTION: SIGNAL PEPTIDES NUCLEIC ACID MOLECULES AND METHODS FOR TREATME
; FILE REFERENCE: P04885US1
; CURRENT APPLICATION NUMBER: US/09/833,017
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/CA00/00605
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 2,302,861
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 2,332,733
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 141
; TYPE: DNA
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(141)
US-09-833-017-1

```
Alignment Scores:
Pred. No.: 1.09e-29 Length: 141
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-833-017b-2 (1-46) x US-09-833-017-1 (1-141)
QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 1 ATGAAAAAACACTAUCATTTAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAGAG 60
QY 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 61 ATTATCATTTGGCGGAGCGGAGCCTATCAACATTTTTCGGCTGTTTAACAGAAAGTTT 120
QY 41 ThrGlnAlaLeuGlyLys 46
Db 121 ACACAAGCTTTGGGAAAA 138

RESULT 2
US-09-833-017-21
; Sequence 21, Application US/09833017
; Patent No. US20020081302A1
; GENERAL INFORMATION:
; APPLICANT: CVITKOVITCH, Dennis
; TITLE OF INVENTION: SIGNAL PEPTIDES NUCLEIC ACID MOLECULES AND METHODS FOR TREATMENT
; FILE REFERENCE: P04865U1
; CURRENT APPLICATION NUMBER: US/09/833,017
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/CA00/00605
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 2,302,861
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 2,332,733
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 2557
; TYPE: DNA
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2557)
US-09-833-017-21
Alignment Scores:
Pred. No.: 5.95e-28 Length: 2557
Score: 227.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-833-017b-2 (1-46) x US-09-833-017-21 (1-2557)
QY 1 MetLysLysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGlu 20
Db 101 ATGAAAAAACACTATCATTTAAAAATGACTTTAAAGAAATTAAGACTGATGAATTAGAG 160
QY 21 IleIleIleGlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 40
Db 161 ATTATCATTTGGCGGAGCGGAGCCTATCAACATTTTTCGGCTGTTTAACAGAAAGTTT 220
QY 41 ThrGlnAlaLeuGlyLys 46
Db 221 ACACAAGCTTTGGGAAAA 238

RESULT 3
```

```
US-09-974-300-4841
; Sequence 4841, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085,500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4841
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-4841
Alignment Scores:
Pred. No.: 0.438 Length: 416
Score: 58.50 Matches: 18
Percent Similarity: 60.53% Conservative: 5
Best Local Similarity: 47.37% Mismatches: 13
Query Match: 25.77% Indels: 2
DB: 10 Gaps: 1

US-09-833-017b-2 (1-46) x US-09-974-300-4841 (1-416)
QY 3 LysThrLeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGluIleI 22
Db 36 AGACATGAGTCTTAAGCGGATTACGACGACATTCGCGAAAGAGTGTGTTGCAT 95
QY 22 eileGlySerGlySer---LeuSerThrPhePheArgLeuPheAsnArg 38
Db 96 TATTGGCCGCTGTGCTCGCGGAAAGTACGTTTCTCCGCTGTTAAACCGG 147

RESULT 4
US-09-974-300-2095/c
; Sequence 2095, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085,500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2095
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-2095
Alignment Scores:
Pred. No.: 2.4 Length: 942
Score: 57.00 Matches: 16
Percent Similarity: 53.49% Conservative: 7
Best Local Similarity: 37.21% Mismatches: 16
Query Match: 25.11% Indels: 4
DB: 10 Gaps: 1
```



```
; SEQ ID NO 5
; LENGTH: 3218
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-788-657-5

Alignment Scores:
Pred. No.: 28 Length: 3218
Score: 55.00 Matches: 15
Percent Similarity: 55.26% Conservative: 6
Best Local Similarity: 39.47% Mismatches: 15
Query Match: 24.23% Indels: 2
DB: 1 Gaps: 1

US-09-833-017B-2 (1-46) x US-09-788-657-5 (1-3218)
Qy 7 LeuLysAsnAppheLys-----GluLeLysThrAspGluLeuGluLeuLeuLeuGly 24
Db 2889 CTACAAATGATATTAAAGATAATTGAGTTGAAATAATGATGAATGTTTCTCTCATAAA 2830
Qy 25 GlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGln 42
Db 2829 GGACTCGGCTCTTTGGTACCGTTGAATTAAGCTTAGACAAGAAATCGTTGCTTGAG 2776

RESULT 13
US-09-070-927A-171
; Sequence 171, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Patrick J. Dillon
; Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESS: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: Hp Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 171:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11871 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 171:
US-09-070-927A-171
```

```
Alignment Scores:
Pred. No.: 206 Length: 11871
Score: 54.50 Matches: 17
Percent Similarity: 53.49% Conservative: 6
Best Local Similarity: 39.53% Mismatches: 15
Query Match: 24.01% Indels: 5
DB: 10 Gaps: 2

US-09-833-017B-2 (1-46) x US-09-070-927A-171 (1-11871)
Qy 1 MetLysLys-----ThrLeuSerLeuLysAsnAppheLysGluLeuLysThr 16
Db 4024 ATGAGAATAATTCGAATGCTACACAGCGATTCGGAATATTTTCAGTAGAGATTGACCAA 4083
Qy 17 AspGluLeuGluLeuLeuLeuLeuGlySerGlySer---LeuSerThrPhePheArgLeu 35
Db 4084 GCGCACTTCGTCTATGTGGTTGGTCCCTCAGGTGCGAGGAAATCAACCTTTATTAGATTA 4143
Qy 36 PheAsnArg 38
Db 4144 ATGTATCGT 4152

RESULT 14
US-09-917-800A-1496
; Sequence 1496, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1496
; LENGTH: 2376
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 X51615
US-09-917-800A-1496

Alignment Scores:
Pred. No.: 27 Length: 2376
Score: 54.00 Matches: 21
Percent Similarity: 43.55% Conservative: 6
Best Local Similarity: 33.87% Mismatches: 9
Query Match: 23.79% Indels: 26
DB: 10 Gaps: 4
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 22:01:57 ; Search time 1808.42 Seconds
(without alignments)
411.958 Million cell updates/sec

Title: US-09-833-017b-2
Perfect score: 227
Sequence: 1 MKKTLKNDKFKIKTDELE.....GSLSTFFRLFNRSFTQALGK 46

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09833017/runat_05112002_105349_4846/app_query.fasta_1.398
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017@cgn.1.1.899 @runat_05112002_105349_4846 -NCPU=6 -ICPU=3
-NO_XLPPX -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECUT=120
-WARN_TIMEOUT=30 -THRGADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_estl.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rpd.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	67	29.5	691	13	BJ069595 BJ069595
2	63	27.8	328	9	AA358436 EST67334
3	62	27.3	527	13	BI707485 fs41f04.y
4	62	27.3	565	10	AW547369 L0021C10-
5	62	27.3	621	13	BI841174
6	62	27.3	720	13	BI891813 fq52b06.y
7	62	27.3	721	12	BI891813 fq52b06.y
8	62	27.3	1142	17	BG587685 EST489460
9	62	27.3	1391	12	AL117217 t7 end of
10	61.5	27.1	661	10	BF207257 601870728
11	61.5	27.1	818	12	AV321480 AV321480
12	61	26.9	250	10	BG678629 602624419
13	61	26.9	906	17	BB067911 BB067911
14	60.5	26.7	560	17	AZ528065 ENTBT050TR
15	60	26.4	483	17	AZ849249 2M0150B07
16	60	26.4	489	12	AF149972 SP_0037.B
17	60	26.4	636	17	BZ435651 bac34f12
18	60	26.4	808	17	BH603995 BQGQ193TF
19	60	26.4	1118	12	BF384061 Tetraodon
20	59.5	26.2	634	10	BB375066 BB375066
21	59	26.0	302	10	AV340411 AV340411
22	59	26.0	394	9	AI553452 v555f10.x
23	59	26.0	506	10	AW861203 KCI-CT030
24	59	26.0	609	10	AW254713 ML100 pep
25	59	26.0	686	14	BU007240 QGH13N12
26	58.5	25.8	436	10	BB819813 BB819813
27	58.5	25.8	624	17	BH385411 AG-ND-146
28	58.5	25.8	661	10	BE201479 fk93d08.x
29	58.5	25.8	789	12	BE877229 601485150
30	58	25.6	432	17	BH770061 BMBAC369F
31	58	25.6	432	17	AQ085216 HS-2270.A
32	58	25.6	609	17	BH404546 AG-ND-137
33	58	25.6	624	9	AL507817 AL507817
34	58	25.6	658	10	BB369772 BB369772
35	58	25.6	744	12	BG597482 EST496160
36	57.5	25.3	263	12	BG055682 naf62h04
37	57.5	25.3	539	10	AV526794 AV526794
38	57.5	25.3	562	13	BG952882 PM4-CT054
39	57.5	25.3	615	12	BG661558 kx03d05.y
40	57.5	25.3	676	10	AV822067 AV822067
41	57.5	25.3	762	17	AQ157280 nbx800090
42	57.5	25.3	869	12	BG105781 602312622
43	57.5	25.3	954	17	AL199640 Tetraodon
44	57	25.1	294	9	AI986937 rs19d12.y
45	57	25.1	333	10	BB095765 BB095765

ALIGNMENTS

RESULT 1
BJ069595
LOCUS
DEFINITION BJ069595 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL053010 5', mRNA sequence.
ACCESSION BJ069595
VERSION BJ069595.1 GI:17497955
KEYWORDS
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 691)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara

please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source

```
1. 250
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="8030468G23"
/clone_lib="RIKEN full-length enriched, 15 days embryo
male testis"
/sex="male"
/tissue_type="testis"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/note="Site.1: Sali; Site.2: BamHI: cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of subtraction to
Rot = 185.0 Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATCTCAGTAAATAATTAATCCGCCCCCCC 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"
```

BASE COUNT

58 a 62 c 41 g 89 t

ORIGIN

Alignment Scores:

```
Pred. No.: 7.8 Length: 250
Score: 61.00 Matches: 12
Percent Similarity: 66.67% Conservative: 6
Best Local Similarity: 44.44% Mismatches: 9
Query Match: 26.87% Indels: 0
DB: 10 Gaps: 0
```

US-09-833-017B-2 (1-46) x BB067911 (1-250)

Qy 2 LysLysThrLeuSerLeuLysAsnAspPhelysGluLeuLysThrAspGluLeuGluLeu 21

Db 117 AGGAAACAATGAAATAAATAATGTGTACAAAGATATAAAGAGGAGGAAATACATA 58

Qy 22 lleleGlySerGlySer 28

Db 57 TGGAGGGTGGGTGGGGAAT 37

RESULT 13

AZ528065/c

LOCUS

DEFINITION ENTBT05TR Entamoeba histolytica sheared DNA Entamoeba histolytica genomic, DNA sequence.

ACCESSION

AZ528065

VERSION

AZ528065.1

KEYWORDS

GSS.

SOURCE

Entamoeba histolytica.

ORGANISM

Entamoeba histolytica

REFERENCE

1 (bases 1 to 906)

AUTHORS

Loftus, B., Van Aken, S. and Fraser, C.

TITLE

Determination of clone end sequences from Entamoeba histolytica

JOURNAL

HMI:IMSS sheared DNA library

COMMENT

Unpublished (2000)

Contact: Brendan J Loftus

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0208

Fax: 301 838 3543

Email: bjloftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared

DNA library

Seq primer: M13-Reverse

Class: shotgun

High quality sequence start: 99

High quality sequence stop: 763.

FEATURES

source

1. 906

/organism="Entamoeba histolytica"

/strain="HMI:IMSS"

/db_xref="taxon:5759"

/clone_lib="Entamoeba histolytica sheared DNA"

/note="Vector: pHOS1; Site.1: Bst I; Constructed at The

Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica

using a method described by Clark and Diamond (Clark,

C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a

method for isolate identification. Exp. Parasitol.

77:450.). The DNA was mechanically sheared to give a

tight size distribution (~2 kb). The v + i method used for

the library construction is described in detail in Smith,

H.O. and Venter, J.C. (Making small insert libraries for

whole genome shotgun sequencing projects. In Genome

Sequencing: A Practical Approach, eds. M. Vaudin and B.

Barell, Oxford University Press, 1999)."

BASE COUNT 311 a 146 c 96 g 353 t

ORIGIN

Alignment Scores:

Pred. No.: 41.2 Length: 906

Score: 61.00 Matches: 14

Percent Similarity: 58.97% Conservative: 9

Best Local Similarity: 35.90% Mismatches: 14

Query Match: 26.87% Indels: 2

DB: 17 Gaps: 1

US-09-833-017B-2 (1-46) x AZ528065 (1-906)

Qy 8 LysAsnAspPhelysGluLeuLysThrAspGluLeuGluLeuLeuLeuGlySerGly 27

Db 730 CAAACGAGCGCAGAGATGTAAAGTAAAAAATTAAGAACATTTGTTGGAGGTAATG 671

Qy 28 SerLeuSerThrPhe-----PheArgLeuPheAsnArgSerPheThrGlnAlaLeu 44

Db 670 GACCATTCACATTGCCCTGTTTAGACTTGATGATACGCTTTTACTGAAAAATA 614

RESULT 14

AZ849249

LOCUS

DEFINITION

clone UUGC2M0150B07 R, DNA sequence.

ACCESSION

AZ849249

VERSION

AZ849249.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 560)

AUTHORS

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.

and Wright, D., Weiss, R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert length: 10000 Std Error: 0.00
Plate: 0150 row: B column: 07
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 560.

FEATURES

Location/Qualifiers
1..560
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUCG2M0150B07"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/note="Vector: pWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWD42 (g14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

source

BASE COUNT 238 a 102 c 70 g 150 t
ORIGIN
Alignment Scores:
Pred. No.: 26.6 Length: 560
Score: 60.50 Matches: 15
Percent Similarity: 55.32% Conservative: 11
Best Local Similarity: 31.91% Mismatches: 18
Query Match: 26.65% Indels: 3
DB: 17 Gaps: 1
US-09-833-017B-2 (1-46) x AZ849249 (1-560)
QY 1 MetLysThrLeuSerLeu-----LysAsnAspPheLysGluIleLysThrAsp 17
Db 112 CTGATGCAACATCTCAATAACACTAGAAAAACAAATTAACAAAAACAAAAAGAG 171
QY 18 GluLeuGluIleIleGlySerGlySerThrPhePheArgLeuPheAsn 37
Db 172 CAATTGCTACTATCATCCAGAGTTCGGGGATGTTTCACATTTTCATCAATCAATCAAT 231
QY 38 ArgSerPheThrGlnAlaLeu 44
Db 232 CAATCAATCAATCAATCAAT 252

RESULT 15

AZ149972

LOCUS

DEFINITION

SP_0037_B1_E08_T7 Strongylocentrotus purpuratus, purple sea urchin,
sperm genomic BAC library Strongylocentrotus purpuratus genomic
clone Plate=37 Col=15 Row=J, DNA sequence.

ACCESSION

AZ149972

VERSION

AZ149972.1 GI:8301873

KEYWORDS

GSS.

SOURCE

Strongylocentrotus purpuratus

ORGANISM

Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;

REFERENCE

AUTHORS

Cameron,R.A., Mahairas,G., Rast,J.P., Martinez,P., Biondi,T.R.,
Swartzell,S., Wallace,J.C., Poustka,A.J., Livingston,B.T., Wray
,G.A., Etensohn,C.A., Lehrach,H., Britten,R.J., Davidson,E.H. and
Hood,L.

TITLE

A sea urchin genome project: Sequence scan, virtual map, and
additional resources

JOURNAL

MEDLINE

COMMENT

Proc. Natl. Acad. Sci. U.S.A. 97 (17), 9514-9518 (2000)

Contact: Cameron, RA, Davidson, EH, Hood, L

Division of Biology 156-29

California Institute of Technology

Pasadena California 91125, USA

Tel: (626) 395-8421

Fax: (626) 793-3047

Email: acameron@caltech.edu

Plate: 37 row: J column: 15

Seq primer: T7

Class: BAC ends

High quality sequence stop: 483.

Location/Qualifiers

1..483

/organism="Strongylocentrotus purpuratus"

/db_xref="taxon:7668"

/clone_lib="plate=37 Col=15,Row=J"

urchin, sperm genomic BAC library"

/note="Organ: sperm; Vector: BACs3.6; BAC Clones in E-Coli
DH10B"

BASE COUNT 151 a 99 c 78 g 142 t 13 others

ORIGIN

Alignment Scores:

Pred. No.: 26.4 Length: 483

Score: 60.00 Matches: 14

Percent Similarity: 62.16% Conservative: 9

Best Local Similarity: 37.84% Mismatches: 12

Query Match: 26.43% Indels: 2

DB: 17 Gaps: 1

US-09-833-017B-2 (1-46) x AZ149972 (1-483)

QY 5 LeuSerLeuLysAsnAspPheLysGluIleLysThrAspGluLeuGluIleIleGly 24

Db 324 ATTGCTCTGAAGAACGATTACAGTAATATTTTCAGATGAGAGATATTCGTCATGGGA 383

QY 25 GlySerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThr 41

Db 384 GGGTGTACATGCGTCCGTTT-----ATCTCGGTGCAAAATAGNACG 428

Search completed: November 8, 2002, 23:39:44

Job time : 1814.42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:02 ; Search time 37.0746 Seconds
(without alignments)
137.814 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 227

Sequence: 1 MKKTLKLNDFKEKTKDELF.....GSLSTFFRLNRSFTQALCK 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	58	25.6	173	AAW28307	Staphylococcus aur
2	56	24.7	939	22 AAU35204	Enterococcus faeca
3	55.5	24.4	553	AAU35204	Sequence of Newcas
4	55.5	24.4	553	AAU35204	Newcastle Disease
5	55.5	24.4	553	AAU35204	Newcastle Disease
6	55.5	24.4	553	AAU35204	Newcastle Disease
7	55.5	24.4	553	AAU35204	Newcastle Disease
8	55.5	24.4	553	AAU35204	Newcastle Disease
9	55.5	24.4	553	AAU35204	Seq ID No: 14 of U
10	55.5	24.4	553	AAU35204	Protein encoded by
11	55.5	24.4	553	AAU35204	Newcastle Disease

12	55	24.2	583	22	AAG82132	S. epidermidis ope
13	54.5	24.0	508	21	AAG44399	Arabidopsis thalia
14	54.5	24.0	515	21	AAG44398	Arabidopsis thalia
15	54	23.8	96	14	AAR35757	Monellin protein.
16	54	23.8	97	14	AAR34192	Monellin protein.
17	54	23.8	2353	17	AAR99393	Haemophilus adhesi
18	54	23.8	2411	21	AAB23860	Haemophilus influe
19	53.5	23.6	553	12	AAR10065	Newcastle disease
20	53	23.3	96	21	AAG57201	Arabidopsis thalia
21	53	23.3	513	15	AAR54214	L.lactis branched
22	52	22.9	148	22	ABG28821	Novel human diagno
23	52	22.9	365	22	AAW78620	Human protein SEQ
24	52	22.9	381	22	AAW79604	Human protein SEQ
25	52	22.9	381	22	AAW79605	Human protein SEQ
26	52	22.9	382	22	AAW78621	Human protein SEQ
27	52	22.9	516	21	AAU52007	M. jannaschii M08
28	52	22.9	516	21	AAU51636	M. jannaschii M08
29	51.5	22.7	177	22	ABG17731	Novel human diagno
30	51.5	22.7	257	22	AAG98329	Escherichia coli p
31	51.5	22.7	306	21	AAU15918	E. coli proliferat
32	51.5	22.7	306	22	AAG98991	E. coli growth and
33	51.5	22.7	591	22	ABG10345	Novel human diagno
34	51.5	22.7	2737	22	ABG18288	Novel human diagno
35	51.5	22.7	3076	22	ABG07038	Staphylococcus aur
36	51	22.5	220	22	AAU37294	Human colon cancer
37	51	22.5	292	22	AAG74939	Mutant uracil DNA
38	51	22.5	313	18	AAW21821	Mutant uracil DNA
39	51	22.5	313	18	AAW21822	Mutant uracil DNA
40	51	22.5	313	18	AAW21824	Mutant uracil DNA
41	51	22.5	313	18	AAW21825	Mutant uracil DNA
42	51	22.5	313	18	AAW21814	Uracil DNA glycosy
43	51	22.5	313	18	AAW21815	Cytosine DNA glyco
44	51	22.5	313	18	AAW21816	Thymine DNA glycos
45	51	22.5	313	18	AAW21819	Mutant uracil DNA

ALIGNMENTS

RESULT 1
AAW28307
ID AAW28307 standard; Protein; 173 AA.
AC AAW28307;
XX DT 01-SEP-1998 (first entry)
XX Staphylococcus aureus protein of unknown function.
DE Staphylococcus aureus protein; ribozyme; antisense sequence; control;
KW Staphylococcal gene; regulatory element; bacterial gene expression;
KW vaccine; Staphylococcal infection; food poisoning; scaled skin syndrome;
KW toxic shock syndrome.
XX OS Staphylococcus aureus.
XX FH Key Location/Qualifiers
XX Misc-difference 123 /note= "X is not defined in the specification"
FT FT Misc-difference 132 /note= "X is not defined in the specification"
FT FT Misc-difference 132 /note= "X is not defined in the specification"
XX

PN WO9730070-A1.
XX 21-AUG-1997.
XX 19-FEB-1997; 97WO-US02318.
XX 20-FEB-1996; 96US-0011888.
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX Black MT, Burnham MK, Hodgson JE, Knowles DJC, Nicholas RO;

PI Pratt JM, Reichard RW, Rosenberg M, Ward JM;
 DR WPI; 1997-424969/39.
 DR N-PSDB; AAT84204.
 XX
 PT Novel polypeptide(s) from Staphylococcus aureus strain WCUH29 - used
 PT to isolate antimicrobial compounds, and in vaccines against S.
 PT aureus infection
 XX
 PS Claim 6; Page 597-598; 989pp; English.
 XX
 CC The present sequence represents a Staphylococcus aureus protein of
 CC unknown function. The DNA sequence was isolated from a library of
 CC clones of S. aureus WCUH 29 in Escherichia coli. The DNA sequence can
 CC be used in the construction of ribozymes and antisense sequences to
 CC control the expression of Staphylococcal genes. The DNA sequence is
 CC also useful as a source of regulatory elements for the control of
 CC bacterial gene expression. The present protein may be used to produce
 CC vaccines to enable a host to produce specific antibodies with
 CC antibacterial action. These vaccines and antibodies would protect
 CC a host against invasion by S. aureus, and conditions relating to
 CC Staphylococcal infection, e.g. Staphylococcal food poisoning, scaled
 CC skin syndrome, and toxic shock syndrome.
 XX
 SQ Sequence 173 AA;
 Query Match 25.6%; Score 58; DB 18; Length 173;
 Best Local Similarity 34.3%; Pred. No. 3.3;
 Matches 12; Conservative 8; Mismatches 15; Indels 0; Gaps 0;
 QY 9 NDFKEIKTELEIIIGSGSLSTFFRLNKSFTQA 43
 DB 86 NDFPELFTDLVNSISANKDITKWFQYINKSLLSA 120
 RESULT 2
 AAU35204
 ID AAU35204 standard; Protein; 939 AA.
 XX
 AC AAU35204;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Enterococcus faecalis cellular proliferation protein #491.
 XX
 KW Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO200170955-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WO-US09180.
 XX
 PR 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX
 DR WPI; 2001-611495/70.
 DR N-PSDB; AAS53063.
 XX
 PT New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -
 Example 3; Seq ID No 10797; 511pp; English.
 The invention relates to antisense inhibitors of genes essential to
 prokaryotic cellular proliferation, their use in identifying the
 genes, their use in the discovery of novel antibiotics, the essential
 genes themselves and the encoded proteins. The prokaryotes used are
 Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 invention is also useful for the identification of potential new targets
 for antibiotic development. The antisense nucleic acids can also be used
 to identify proteins used in proliferation, to express these proteins,
 and to obtain antibodies capable of binding to the expressed proteins.
 The proteins can be used to screen compounds in rational drug discovery
 programmes. The antisense nucleic acid sequence is also useful to screen
 for homologous nucleic acids which are required for cell proliferation in
 a wide variety of organisms. The present sequence represents an
 essential prokaryotic cellular proliferation protein.
 Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic
 format directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 939 AA;
 Query Match 24.7%; Score 56; DB 22; Length 939;
 Best Local Similarity 33.3%; Pred. No. 45;
 Matches 11; Conservative 7; Mismatches 15; Indels 0; Gaps 0;
 QY 5 LSLKNDPKEIKTELEIIIGSGSLSTFFRLFN 37
 DB 330 IDMTPEELPADQOEIILNGSGEKNFHYEN 362
 RESULT 3
 AAP70176
 ID AAP70176 standard; protein; 553 AA.
 XX
 AC AAP70176;
 XX
 DT 03-APR-1991 (first entry)
 XX
 DE Sequence of Newcastle Disease Virus (NDV) F polypeptide.
 XX
 KW Epitope; probe; diagnosis.
 XX
 OS Newcastle Disease Virus.
 XX
 PN EP227414-A.
 XX
 PD 01-JUL-1987.
 XX
 PF 16-DEC-1986; 86EP-0309804.
 XX
 PR 15-JUL-1986; 86US-0885765.
 PR 18-DEC-1985; 85GB-0031147.
 PR 14-APR-1986; 86GB-0009037.
 XX
 PA (NATR) NATIONAL RES DEV CORP.
 XX
 PI Bingham RW, Chambers P, Emmerson PT, Millar NS;
 XX
 DR WPI; 1987-179630/26.
 DR N-PSDB; AAN70261.
 XX
 XX Newcastle disease virus gene clones - comprise polynucleotide(s)
 PT encoding the HN and/or F protein of Newcastle disease virus RNA
 PT
 PS Example; pages 11-16; 22pp; English.
 XX
 CC An artificial polynucleotide encoding an HN and/or F polypeptide of
 CC Newcastle Disease Virus (NDV) DNA, a bioprecursor of the

[illegible]

PS Disclosure; Column 61-70; 108pp; English.

XX The invention relates to a recombinant fowlpox virus (FPV) comprising

CC a foreign DNA inserted into a region of the fowlpox virus genome

CC corresponding to a 2.8 kb EcoRI fragment, capable of being expressed in a

CC host cell. The virus is used as a vaccine for immunising chickens against

CC Newcastle disease virus (NDV), Fowlpox, and infectious laryngotracheitis.

XX SQ Sequence 553 AA;

Query Match 24.4%; Score 55.5; DB 20; Length 553;

Best Local Similarity 33.3%; Pred. No. 29;

Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;

OY 4 TSLKNDP-----KEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46

DB 433 TRLSGEDVTYQKNISQDSQVITGNDLIDISTELGNVNSISNALNK 480

RESULT 9

AAB36039

ID AAB36039 standard; Protein; 553 AA.

XX AAB36039;

XX 02-MAR-2001 (first entry)

XX Protein encoded by NDV SfII fragment.

XX Fowlpox virus; FPV; antiviral; antibacterial; vaccine;

KW Newcastle's disease virus; NDV; Marek's disease;

KW Infectious laryngotracheitis.

XX Newcastle's disease virus.

OS US6136318-A.

XX 24-OCT-2000.

XX 07-JUN-1995; 95US-0486414.

XX 26-FEB-1993; 93US-0024156.

PR 28-FEB-1994; 94WO-US02252.

XX (JUNK/) JUNKER D E.

PA (COCH/) COCHRAN M D.

XX Cochran MD, Junker DE;

PI WPI; 2000-686071/67.

XX N-PSDB; AAC67862.

XX New recombinant fowlpox virus useful as vaccines contains foreign DNA

PT inserted into specific non-essential region of the genome

XX Disclosure; Column 71-74; 56pp; English.

XX The present sequence is provided in a specification relating to a

CC recombinant fowlpox virus (FPV) that comprises a foreign DNA inserted

CC within a 3.5 kb EcoRI fragment of the FPV genomic DNA. The foreign DNA

CC can be expressed in host cells infected with FPV. The recombinant FPV

CC may be used in vaccines to protect animals (especially chickens) against

CC fowlpox and, depending on the source of the foreign DNA, other diseases,

CC particularly Newcastle's disease, Marek's disease or infectious

CC laryngotracheitis.

XX SQ Sequence 553 AA;

Query Match 24.4%; Score 55.5; DB 21; Length 553;

Best Local Similarity 33.3%; Pred. No. 29;

Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;

OY 4 TSLKNDP-----KEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46

DB 433 TRLSGEDVTYQKNISQDSQVITGNDLIDISTELGNVNSISNALNK 480

RESULT 10

AAY51231

ID AAY51231 standard; Protein; 553 AA.

XX AAY51231;

XX 07-APR-2000 (first entry)

XX Newcastle disease virus LaSota genome encoded protein 4.

XX Avian-paramyxovirus; infection; lentogenic; F protein; vaccine;

KW respiratory disease; gastrointestinal disease; poultry pathogen;

KW local immunity.

XX Newcastle disease virus.

OS WO9966045-A1.

XX 23-DEC-1999.

XX 17-JUN-1999; 99WO-NL00377.

XX 19-JUN-1998; 98EP-0202054.

XX (DIEN-) STICHTING DIENST LANDBOUWKUNDIG ONDERZOE.

XX Peeters BPH, De Leeuw OS, Koch G, Gielkens ALJ;

XX WPI; 2000-106102/09.

XX New avian paramyxovirus cDNA, useful for production of vaccine against

PT Newcastle disease virus

XX Disclosure; Fig 3; 115pp; English.

XX This invention describes a novel avian-paramyxovirus cDNA (I) which

CC comprises a nucleic acid sequence corresponding to the 5' terminal

CC end of the genome of avian-paramyxovirus allowing the generation of

CC an infectious copy of avian-paramyxovirus. The cell line is useful for

CC the production of infectious lentogenic NDV (Newcastle Disease virus)

CC without the addition of exogenous proteolytic activity. Also it is

CC possible to generate a stable transfected cell line that expresses the

CC wild-type F protein in the virus envelope therefore providing infectious

CC particles, useful in the form of a vaccine, especially against

CC respiratory and/or gastrointestinal diseases. NDV can be easily cultured

CC to very high titers in embryonated eggs. Mass culture of embryonated

CC eggs is relatively cheap. NDV vaccines are relatively stable and can be

CC simply administered by mass application methods e.g. drinking water or

CC by spraying or by aerosol formation. The natural route of infection is

CC by the respiratory and/or gastrointestinal tract which are also the major

CC routes of infection of many other poultry pathogens. NDV can induce local

CC immunity despite the presence of circulating maternal antibody. This

CC sequence represents a protein encoded by the NDV strain LaSota genome

CC which is described in the method of the invention.

XX SQ Sequence 553 AA;

Query Match 24.4%; Score 55.5; DB 21; Length 553;

Best Local Similarity 33.3%; Pred. No. 29;

Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;

OY 4 TSLKNDP-----KEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46

DB 433 TRLSGEDVTYQKNISQDSQVITGNDLIDISTELGNVNSISNALNK 480

RESULT 11

AAY58183

ID AAY58183 standard; Protein; 553 AA.


```
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.0%; Score 54.5; DB 21; Length 508;
Best Local Similarity 39.6%; Pred. No. 36;
Matches 21; Conservative 8; Mismatches 11; Indels 13; Gaps 5;

QY 1 MKKTLKND-FKE-----IKTDELEIII-GSGSLS---TF---PRLENRGF 40
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MEKSTLKNDSFVQYNPVETGSLSIIVLGASDDLAKKKTTPALNLFHQGF 53

RESULT 14
AAG44398
ID AAG44398 standard; Protein; 515 AA.
XX
AC AAG44398;
XX
DT 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55611.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW
```

```
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 18-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
```

```
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0145921.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147132.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 24.0%; Score 54.5; DB 21; Length 515;
Best Local Similarity 39.6%; Pred. NO. 36;
Matches 21; Conservative 8; Mismatches 11; Indels 13; Gaps 5;

QY 1 MKKTLKND-FKE----IKTDELEII--GGSGSL--TF--FRLNRSF 40
DB 8 MEKRTLNDSFVOEYNPVTGSLSIIVLGASGDLAKKTFPALNLFHQGF 60

RESULT 15
AAR35757
ID AAR35757 standard; Protein; 96 AA.
XX
AC AAR35757;
XX
DT 29-JUL-1993 (first entry)
XX
DE Monellin protein.
XX
KW Sweet; heat; resistance; protease.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 51 /label= GLY, SER, GLU
XX
PN JP05070494-A.
```


RECORD 3
US-08-023-610-13
; Sequence 13, Application US/08023610
; Patent No. 5928648
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Macdonald Ph.D., Richard D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/023.610
; FILING DATE: February 26, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

; NAME: White Esq. John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-023-610-13
;
Query Match 24.4%; Score 55.5; DB 2; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels
;
QY 4 TLSLKDNF-----KEIKTDELEIIIGGSGLSITFFRFLNRSFTQALGK 46
||| || | | : :|| | : || | | : || |
DB 433 TRLSGEFDVYQKNISIQDSQVIITNLDISTELGNVNNISNALNK 480
;
RESULT 4
US-08-288-065A-13
; Sequence 13, Application US/08288065A
; Patent No. 5961982
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Macdonald, Richard D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-
; TITLE OF INVENTION: HVT-050 and Uses Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,065A
; FILING DATE: Aug-09-94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-288-065A-13
;
Query Match 24.4%; Score 55.5; DB 2; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels
;
QY 4 TLSLKDNF-----KEIKTDELEIIIGGSGLSITFFRFLNRSFTQALGK 46
||| || | | : :|| | : || | | : || |
DB 433 TRLSGEFDVYQKNISIQDSQVIITNLDISTELGNVNNISNALNK 480
;
RESULT 5
US-08-362-240A-13

```

; Sequence 13, Application US/08362240A
; Patent No. 5965138
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Junker, David
; APPLICANT: Wild, Martha A
; TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/362,240A
; FILING DATE: Dec-22-94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-362-240A-13

Query Match 24.4%; Score 55.5; DB 2; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;
QY 4 TSLKKNDF-----KEIKTDELEIIIGSGSLSTFFFLNRSFTQALGK 46
DB 433 TRLSGEFDVYQKNISIQDSQVITGNLDISTELGNVNNISNALNK 480

RESULT 6
US-08-477-459-14
; Sequence 14, Application US/08477459
; Patent No. 6001369
; GENERAL INFORMATION:
; APPLICANT: Mark D. Cochran
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,459
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-477-459-14
Query Match 24.4%; Score 55.5; DB 3; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;
QY 4 TSLKKNDF-----KEIKTDELEIIIGSGSLSTFFFLNRSFTQALGK 46
DB 433 TRLSGEFDVYQKNISIQDSQVITGNLDISTELGNVNNISNALNK 480
RESULT 7
US-08-479-869-14
; Sequence 14, Application US/08479869
; Patent No. 6123949
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D, Mark D
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,869
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/024,156
; FILING DATE: 26-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-479-869-14

Query Match 24.4%; Score 55.5; DB 3; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;
QY 4 TSLKKNDF-----KEIKTDELEIIIGSGSLSTFFFLNRSFTQALGK 46
DB 433 TRLSGEFDVYQKNISIQDSQVITGNLDISTELGNVNNISNALNK 480

Query Match 24.4%; Score 55.5; DB 4; Length 553;
Best Local Similarity 33.3%; Pred. No. 3.4;
Matches 16; Conservative 7; Mismatches 20; Indels 5; Gaps 1;

```

RESULT 11
PCT-US94-02252A-14
; Sequence 14, Application PC/TUS9402252A
;
; GENERAL INFORMATION:
;
; APPLICANT: Syntro Corporation, et al.
;
; TITLE OF INVENTION: Recombinant Fowlpox Viruses and Uses Thereof
;
; NUMBER OF SEQUENCES: 20
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: John P. White
;
; STREET: 30 Rockefeller Plaza
;
; CITY: New York
;
; STATE: New York
;
; COUNTRY: USA
;
; ZIP: 10112
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
;
; APPLICATION NUMBER: PCT/JUS94/02252A

```


RESULT 13

```
; PRIOR APPLICATION NUMBER: PCT/JP97/02924
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: JP 8/241062
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Candida utilis
US-09-242-690A-6

Query Match      23.8%; Score 54; DB 4; Length 97;
Best Local Similarity 41.4%; Pred. No. 0.59;
Matches 12; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

Qy      1 MKKTLKNDPFEIKTDELEIIIGGSL 29
Db      43 MKKIYEENGFEIKGYEYQLVYASDKL 71

Search completed: November 5, 2002, 10:58:19
Job time : 16.4179 secs
```


ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
T.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: AB9758; PMID:21311952; PMID:11418146
A:Accession: E89792
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1009 <KUR>
A:Cross-references: GB:BA000018; PID:g13700198; PIDN:BAB41496.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0272

Query Match 25.6%; Score 58; DB 2; Length 1009;
Best Local Similarity 34.3%; Pred. No. 34;
Matches 12; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Qy 9 NDFEIKTDLEIIIGSGSLSTFFRLFNRSFTQA 43
||| ||| | : : : ||| | -
Db 202 NDFPELFTDLVNSISANKDITKWFTYINKLSLSA 236

RESULT 6
T52611
glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49) [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 03-Nov-2000
C:Accession: T52611
R:Wendt, U.K.; Haenschild, R.; Lange, C.; Pietersma, M.; Wenderoth, I.; von Schaeuwen,
Plant Mol. Biol. 40, 487-494, 1999
A:Title: Evidence for functional convergence of redox regulation in G6PDH isoforms of
A:Reference number: Z25284
A:Accession: T52611
A>Status: preliminary; translated from GB/EMBL/DDBB
A:Molecule type: mRNA
A:Residues: 1-516 <WEN>
A:Cross-references: EMBL:AJO10970; PIDN:CAB52674.1
C:Genetics:
A:Gene: acg9
C:Superfamily: glucose-6-phosphate dehydrogenase
C:Keywords: oxidoreductase

Query Match 25.3%; Score 57.5; DB 2; Length 516;
Best Local Similarity 41.5%; Pred. No. 19;
Matches 22; Conservative 7; Mismatches 11; Indels 13; Gaps 5;

Qy 1 MKKTLSLND--FRP-----IKTDELEIII-CGSGSLS---TF---FRLFNRSF 40
| | : ||| | | : | | | : | | | : | | : | |
Db 8 MEKRSTLNDSFVKEPYNPTGTSLIIVLGASGDLAKKKTFPALNFLHQGF 60

RESULT 7
AC3578
glycine betaine/l-proline transport ATP-binding protein prov BMEI0548 [imported] - Brucella melitensis
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AC3578
R:DelVecchio, V.G.; Kapatalai, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.
A:Reference number: AD3252; PMID:11756688
A:Accession: AC3578
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-398 <KUP>
A:Cross-references: GB:AE008918; PIDN:AAL53790.1; PID:g17984720; GSPDB:GN00191
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI0548
A:Map position: II

A:Gene: SGD:583502z
 A:Cross-references: SGD:S0004082; MIPS:YL092W |
 A:Map position: 12R
 C:Keywords: transmembrane protein
 F:136-152/Domain: transmembrane #status predicted <TM1>
 F:223-239/Domain: transmembrane #status predicted <TM2>
 F:244-260/Domain: transmembrane #status predicted <TM3>
 F:302-318/Domain: transmembrane #status predicted <TM4>
 F:358-374/Domain: transmembrane #status predicted <TM5>
 F:411-427/Domain: transmembrane #status predicted <TM6>
 F:488-504/Domain: transmembrane #status predicted <TM7>
 F:538-554/Domain: transmembrane #status predicted <TM8>
 F:557-573/Domain: transmembrane #status predicted <TM9>

```

RESULT 11
T12258
pyrroline-5-carboxylate synthetase (EC 1.5.1.1.) - common ice plant
C:Species: Mesembryanthemum crystallinum (common ice plant)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000
C:Accession: T12258

```

QY 6 SLANDKEIKTDELEIIIGSGSLST-----FRLNRSF 40
 || ||: :| ||: | ||: | ||
Db 37 SLCEQLKELNSDGYEILVTSGAVSAGRQLRFRKLNSSF 77

```

RESULT 12
E81339
probable restriction /modification enzyme Cj0690c [imported] - Campylobacter jejuni
C.Species: Campylobacter jejuni
C.Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Mar-2000
C.Accession: E81339
R.Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chli
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A.Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A.Reference number: AB1250; MUID:20150912
A.Accession: E81339
A.Status: preliminary

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:03 ; Search time 10.9851 Seconds
(without alignments)
162.138 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 227

Sequence: 1 MKKTLKNDKFEIKTDELE.....GSLSTFFRLFNRSFTQALGK 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60.5	26.7	475	1 YMT4_YEAST	Q04213 saccharomyc
2	58.5	25.8	41	1 CSPL_STRPN	Q54712 streptococ
3	56.5	24.9	893	1 SUL2_YEAST	Q12325 saccharomyc
4	56	24.7	719	1 P5CS_MESCR	O65361 m delta 1-p
5	55.5	24.4	41	1 CSP_STROR	O33689 streptococ
6	55.5	24.4	553	1 VGLF_NDVH	P06156 newcastl d
7	55.5	24.4	553	1 VGLF_NDVH4	P33613 newcastl d
8	55.5	24.4	553	1 VGLF_NDVH	P33614 newcastl d
9	55.5	24.4	553	1 VGLF_NDVH	P12571 newcastl d
10	55	24.2	608	1 CB32_YEAST	P40969 saccharomyc
11	54.5	24.0	373	1 HIS8_METJA	Q58365 methanococ
12	54.5	24.0	463	1 DNAA_RICCN	O92b56 rickettsia
13	54.5	24.0	720	1 SYFB_PORPU	P51346 porphyra pu
14	54	23.8	226	1 CXB2_MOUSE	Q00977 mus musculu
15	54	23.8	226	1 CXB2_RAT	P21994 rattus norv
16	54	23.8	708	1 YNZB_CABEL	P45972 caenorhabd
17	54	23.8	1292	1 PPOC_MYCGE	P47582 mycoplasma
18	53.5	23.6	259	1 PSTB_EDWTA	Q9am14 edwardsiell
19	53.5	23.6	553	1 VGLF_NDVH	P35936 newcastl d
20	53.5	23.6	553	1 VGLF_NDVH	P33615 newcastl d
21	53.5	23.6	553	1 VGLF_NDVH	P12570 newcastl d
22	53.5	23.6	617	1 CYG2_HUMAN	O75343 homo sapien
23	53	23.3	274	1 PNPK_AQUAE	O67055 aquifex aeo
24	53	23.3	400	1 PROV_ECOLI	P14175 escherichia
25	53	23.3	400	1 PROV_SALTY	P17328 lactomonella
26	53	23.3	513	1 LEUL_LACLA	Q02141 lactococcus
27	53	23.3	591	1 YN48_YEAST	P28466 saccharomyc
28	53	23.3	1132	1 DNBL_HSV6U	P52338 human herpe
29	53	23.3	1132	1 DNBL_HSV6Z	P52338 human herpe
30	52.5	23.1	310	1 GTRB_BP22	P57022 bacterioph
31	52.5	23.1	564	1 Y61A_MYCPN	P75041 mycoplasma
32	51.5	22.7	41	1 CSP2_STROR	O33690 streptococ
33	51.5	22.7	240	1 PYRH_METJA	Q58656 methanococ

ALIGNMENTS

```

RESULT 1
YMT4_YEAST          STANDARD;          PRT;   475 AA.
ID YMT4_YEAST
AC Q04213;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Hypothetical 55.4 kDa protein in MCM1-NUP116 intergenic region.
GN YMR044W OR YN9532.09.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Odell C., Bowman S., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 248502; CAA88410.1; -
DR SGD; S0004647; YMR044W.
DR InterPro; IPR000313; PWMF.
DR Pfam; PF00855; PWMF; 1.
DR SMART; SM00293; PWMF; 1.
KW Hypothetical protein.
FT DOMAIN 74 77 POLY-GLU.
FT DOMAIN 186 195 POLY-GLU.
FT DOMAIN 243 251 POLY-GLU.
SQ SEQUENCE 475 AA; 55426 MW; DFD6189B81178060 CRC64;

Query Match 26.7%; Score 60.5; DB 1; Length 475;
Best Local Similarity 38.9%; Pred. No. 2.8;
Matches 14; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

OY 1 MKKTLKNDKFEIKTDELEIIIGSGSLSTFFRLF 36
   : : : : : : : : : : : : : : : : : : : :
Db 361 INQAVSIKDDF-EIILDELQIALDTRGRNEFITF 395

RESULT 2
CSPL_STRPN          STANDARD;          PRT;   41 AA.
ID CSPL_STRPN
AC Q54712;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Competence stimulating peptide type 1 precursor (CSP-1).
GN COMC1 OR COMC.

```



```

CC -----
DR EMBL; A5008645; AAL03454.1; -.
DR InterPro; IPR001957; Bac_DnaA.
DR Pfam; PF00308; bac_dnaA; 1.
DR PROSITE; PS01008; DNAA; 1.
KW DNA replication; DNA-binding; ATP-binding; Complete proteome.
FT NP_BIND 168 175 ATP (POTENTIAL).
SQ SEQUENCE 463 AA; 52942 MW; B48D2FE43A70A8EA CRC64;
Query Match 24.0%; Score 54.5; DB 1; Length 463;
Best Local Similarity 28.2%; Pred. No. 16;
Matches 11; Conservative 8; Mismatches 17; Indels 3; Gaps 1;
QY 2 KPTLSLKDFKEIKT---DELEITIGSGSLSTFRLEN 37
DB 213 KEVISFKKFRSDVLMIDDIQICGKSTQEEFFHTFN 251

RESULT 13
SYFB_PORPU STANDARD; PRT; 720 AA.
AC SFYFB_PORPU
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--
DE tRNA ligase beta chain) (PHERS).
GN PHET.
OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVONPORT;
RA Reith M.E., Munnolland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
RL genome.";
CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +
CC diphosphate + L-phenylalanyl-tRNA(Phe).
CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA
CC CHAIN FAMILY. SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38804; AAC08232.1; -.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Chloroplast.
SQ SEQUENCE 720 AA; 82592 MW; 495E49947117F62A CRC64;
Query Match 24.0%; Score 54.5; DB 1; Length 720;
Best Local Similarity 31.5%; Pred. No. 26;
Matches 17; Conservative 7; Mismatches 9; Indels 21; Gaps 2;
QY 6 SLANDFKETKTDELEITIGSGS-----GSLSTFRLENRS 39
DB 495 NLARN-KIETTKLAILGSLDIRSEWSEPAHSLNWEAKGLIENFFFKLNKS 547

RESULT 14
CXB2_MOUSE STANDARD; PRT; 226 AA.
ID CXB2_MOUSE
AC Q00977;

```

```

DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction beta-2 protein (Connexin 26) (Cx26).
GN GJB2 OR CXN-26.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=92354576; PubMed=1322820;
RA Willecke K., Nicholson B.J., Dahl E., Kozjek G., Hennemann H.;
RT "Molecular cloning of mouse connexins26 and -32: similar genomic
RT organization but distinct promoter sequences of two gap junction
RT genes.";
RL Eur. J. Cell Biol. 58:81-89(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=91285228; PubMed=2060697;
RA Nishi M., Kumar N.M., Gilula N.B.;
RT "Developmental regulation of gap junction gene expression during
RT mouse embryonic development.";
RL Dev. Biol. 146:117-130(1991).
CC -1- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -1- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, INTESTINE, LUNG, SPLEEN,
CC STOMACH, TESTIS, & BRAIN, BUT NOT HEART & ADULT SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP I)
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81445; AAA37495.1; -.
DR EMBL; M63803; AAA37276.1; -.
DR PIR; C49769; C49769.
DR MG; MG1:95720; G3b2.
DR InterPro; IPR000500; Connexin.
DR Pfam; PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR SMART; SM00037; CNX; 1.
DR PROSITE; PS00407; CONNEXINS_1; 1.
DR PROSITE; PS00408; CONNEXINS_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 20 40 POTENTIAL.
FT DOMAIN 41 75 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 76 96 POTENTIAL.
FT DOMAIN 97 143 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 164 POTENTIAL.
FT DOMAIN 165 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 POTENTIAL.
FT DOMAIN 211 226 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 68 68 H -> Y (IN REF. 2).
FT CONFLICT 215 215 V -> I (IN REF. 2).
SQ SEQUENCE 226 AA; 26411 MW; OD8178F7A339E0B6 CRC64;
Query Match 23.8%; Score 54; DB 1; Length 226;
Best Local Similarity 33.9%; Pred. No. 8.5;
Matches 21; Conservative 6; Mismatches 9; Indels 26; Gaps 4;
QY 7 LKNDFK---EIKTDELEITIGSGSL-----STFR-----LFNRSFTQAL 44
DB 111 IKNEFKDIEIKTKVRI-----EGSLWVYTTTSIFRFRVFEAFVYFYIMNGFFMQRL 166

```

Search completed: November 5, 2002, 10:57:52
Job time : 13.9851 secs

QY 45 GK 46
|
Db 167 VK 168

RESULT 15

CXB2_RAT STANDARD; PRT: 226 AA.
AC P21994;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gap junction beta-2 protein (Connexin 26) (Cx26).
GN GJB2 OR CXN-26.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=90094541; PubMed=2557354;
RA Zhang J.T., Nicholson B.J.;
RT "Sequence and tissue distribution of a second protein of hepatic gap
RT junctions, Cx26, as deduced from its cDNA.";
RL J. Cell Biol. 109:3391-3401(1989).
RN [2]
RP SEQUENCE OF 1-18.
RX MEDLINE=88039049; PubMed=2823143;
RA Nicholson B.J., Dermietzel R., Teplow D., Traub O., Willecke K.,
RA Revel J.-P.;
RT "Two homologous protein components of hepatic gap junctions.";
RL Nature 329:732-734(1987).
CC -|- FUNCTION: ONE GAP JUNCTION CONSISTS OF A CLUSTER OF CLOSELY PACKED
CC PAIRS OF TRANSMEMBRANE CHANNELS, THE CONNEXONS, THROUGH WHICH
CC MATERIALS OF LOW MW DIFFUSE FROM ONE CELL TO A NEIGHBORING CELL.
CC -|- SUBUNIT: A CONNEXON IS COMPOSED OF A HEXAMER OF CONNEXINS.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- TISSUE SPECIFICITY: LIVER, KIDNEY, INTESTINE, LUNG, SPLEEN,
CC STOMACH, TESTIS, & BRAIN, BUT NOT HEART & ADULT SKELETAL MUSCLE.
CC -|- SIMILARITY: BELONGS TO THE CONNEXIN FAMILY. BETA-TYPE (GROUP 1)
CC SUBFAMILY.
CC PIR; A33646; A33646.
DR InterPro: IPR000500; Connexin.
DR Pfam: PF00029; connexin; 1.
DR PRINTS; PR00206; CONNEXIN.
DR SMART; SM00037; CNX; 1.
DR PROSITE; PS00407; CONNEXINS_1; 1.
DR PROSITE; PS00408; CONNEXINS_2; 1.
KW Gap junction; Transmembrane.
FT DOMAIN 1 19 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 20 40 POTENTIAL.
FT DOMAIN 41 75 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 76 96 POTENTIAL.
FT DOMAIN 97 143 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 144 164 POTENTIAL.
FT DOMAIN 165 189 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 190 210 POTENTIAL.
FT DOMAIN 211 226 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 226 AA; 26451 MW; E36EC95F1B235A0D CRC64;

Query Match 23.8%; Score 54; DB 1; Length 226;
Best Local Similarity 33.9%; Pred. No. 8.5;
Matches 21; Conservative 6; Mismatches 9; Indels 26; Gaps 4;

QY 7 LKNDPK---EKTDLEIIIGSGSL-----STEF-----LFNRSFTQAL 44
:||||| :||| :| ||| |||||
Db 111 IKNEFKDIEIKTKVKRI----EGSUMWYTTSTFFRFEAFVFMVYFMINGFFMQRL 166
:||||| :||| :| ||| |||||

QY 45 GK 46
|
Db 167 VK 168

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:02 ; Search time 29.5224 Seconds
(without alignments)
269.550 Million cell updates/sec

Title: US-09-833-017b-2

Perfect score: 227

Sequence: 1 MKKTLKNDFKIKTKDELE.....GSLSTFFRLNRSFTQALGK 46

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL19.*

- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-rodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-virus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	227	100.0	46	2 Q99QI5	Q99QI5 streptococc
2	220	96.9	46	2 Q9APK7	Q9APK7 streptococc
3	209	92.1	43	2 Q9APK6	Q9APK6 streptococc
4	69	30.4	850	16 Q97N40	Q97N40 streptococc
5	63.5	28.0	648	2 Q9L651	Q9L651 lactococcus
6	59.5	26.2	275	16 Q9CPE8	Q9CPE8 pasteurilla
7	58	25.6	48	2 Q33596	Q33596 streptococc
8	58	25.6	1009	16 Q99WU3	Q99WU3 staphylococ
9	57.5	25.3	453	3 Q9CIA3	Q9CIA3 saccharomyc
10	57.5	25.3	516	10 Q9SUK0	Q9SUK0 arabidopsis
11	57.5	25.3	516	10 Q9LK23	Q9LK23 arabidopsis
12	57	25.1	591	2 Q9AHY9	Q9AHY9 photorhabdu
13	56.5	24.9	407	2 Q9RQ06	Q9RQ06 lactococcus
14	56.5	24.9	407	2 Q9L931	Q9L931 lactococcus
15	56.5	24.9	408	16 Q9KIF7	Q9KIF7 lactococcus
16	56.5	24.9	593	16 Q92GD6	Q92GD6 rickettsia

17	56	24.7	1250	16 Q9PPL7	Q9PPL7 campylobact
18	55.5	24.4	125	12 Q41368	Q41368 newcastle d
19	55.5	24.4	125	12 Q41472	Q41472 newcastle d
20	55.5	24.4	416	16 Q92EB0	Q92EB0 listeria in
21	55.5	24.4	553	12 Q90339	Q90339 newcastle d
22	55.5	24.4	553	12 Q9DL04	Q9DL04 newcastle d
23	55.5	24.4	553	12 Q9NMH7	Q9NMH7 newcastle d
24	55.5	24.4	553	12 Q92150	Q92150 newcastle d
25	55.5	24.4	553	12 Q9WLE2	Q9WLE2 newcastle d
26	55.5	24.4	553	12 Q9WLE1	Q9WLE1 newcastle d
27	55.5	24.4	553	12 Q91HX4	Q91HX4 newcastle d
28	55.5	24.4	553	12 Q91AH8	Q91AH8 newcastle d
29	55.5	24.4	837	10 Q9M819	Q9M819 arabidopsis
30	55.5	24.4	1304	5 Q9U0H4	Q9U0H4 streptococc
31	55	24.2	48	2 Q33615	Q33615 streptococc
32	54.5	24.0	230	16 Q9CGX0	Q9CGX0 lactococcus
33	54.5	24.0	255	2 Q9AKI3	Q9AKI3 rickettsia
34	54.5	24.0	255	2 Q9AKN6	Q9AKN6 rickettsia
35	54.5	24.0	398	16 Q9ALN0	Q9ALN0 streptococc
36	54.5	24.0	463	16 Q92H56	Q92H56 rickettsia
37	54.5	24.0	819	5 Q97330	Q97330 plasmodium
38	54	23.8	48	2 Q33651	Q33651 streptococc
39	54	23.8	235	2 Q9X9J9	Q9X9J9 vibrio para
40	54	23.8	243	16 Q97M22	Q97M22 clostridium
41	54	23.8	2353	2 P71401	P71401 haemophilus
42	53.5	23.6	125	12 Q41386	Q41386 newcastle d
43	53.5	23.6	125	12 Q41388	Q41388 newcastle d
44	53.5	23.6	165	1 Q90392	Q90392 methanococc
45	53.5	23.6	176	2 Q9RJ74	Q9RJ74 streptomyce

ALIGNMENTS

RESULT 1

ID	Q99QI5	PRELIMINARY:	PRT:	46 AA.
AC	Q99QI5			
DT	01-JUN-2001 (TREMREL. 17, Created)			
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)			
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)			
DE	COMPETENCE STIMULATING PROTEIN PRECURSOR.			
GN	COMC.			
OS	Streptococcus mutans.			
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=1309;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN=GB14, H7, LT11, NG8, AND UAL59;			
RX	MEDLINE=21142515; PubMed=11208787;			
RA	Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Vitkovitch D.G.;			
RT	"Natural Genetic Transformation of Streptococcus mutans Growing in Biofilms."			
RL	J. Bacteriol. 183:897-908(2001).			
DR	EMBL; AF277152; RAK01542.1; -			
DR	EMBL; AF277153; RAK01543.1; -			
DR	EMBL; AF277155; RAK01545.1; -			
DR	EMBL; AF277156; RAK01546.1; -			
DR	EMBL; AF277157; RAK01547.1; -			
DR	InterPro; IPR004288; COMC.			
DR	Pfam; PF03047; COMC; 1.			
FT	CHAIN 26 46			
SQ	SEQUENCE 46 AA; 5211 MW; 38FA62B6F78FC3BF CRC64;			

Query Match 100.0%; Score 227; DB 2; Length 46;

Best Local Similarity 100.0%; Pred. No. 5.3e-22;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTLKNDFKIKTKDELEIIIGSGSLSTFFRLNRSFTQALGK 46

Db 1 MKKTLKNDFKIKTKDELEIIIGSGSLSTFFRLNRSFTQALGK 46

```

RESULT 2
Q9APK7
ID Q9APK7 PRELIMINARY; PRT; 46 AA.
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM71;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in
RT Biofilms.";
RL J. Bacteriol. 183:897-908(2001).
DR EMBL: AF277151; AAK01541.1;
FT CHAIN 26 46 COMPETENCE STIMULATING PROTEIN.
SQ SEQUENCE 46 AA; 5195 MW; 38E0B9D5B8FC3BF CRC64;

Query Match 96.9%; Score 220; DB 2; Length 46;
Best Local Similarity 97.8%; Pred. No. 4.2e-21;
Matches 45; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKKTLKNDKFEIKTDELEIIIGSGSLSTFFRLNRSFTQALGK 46
Db 1 MKKTLKNDKFEIKTDELEIIIGSGSLSTFFRLNRSFTQALGK 46

RESULT 3
Q9APK6
ID Q9APK6 PRELIMINARY; PRT; 43 AA.
AC Q9APK6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JHI005;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in
RT Biofilms.";
RL J. Bacteriol. 183:897-908(2001).
DR EMBL: AF277154; AAK01544.1;
FT CHAIN 26 43 COMPETENCE STIMULATING PROTEIN.
SQ SEQUENCE 43 AA; 4927 MW; E6A78FC3BF6156C7 CRC64;

Query Match 92.1%; Score 209; DB 2; Length 43;
Best Local Similarity 97.7%; Pred. No. 1e-19;
Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKKTLKNDKFEIKTDELEIIIGSGSLSTFFRLNRSFTQ 43
Db 1 MKKTLKNDKFEIKTDELEIIIGSGSLSTFFRLNRSFTQ 43

RESULT 4
Q97N40
ID Q97N40 PRELIMINARY; PRT; 850 AA.
AC Q97N40;

```

```

DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ABC TRANSPORTER, PERMEASE PROTEIN, PUTATIVE.
GN SP2231.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D., C.L.,
RA Holtzapple E., Khouiri H., Wolf A.M., Utterback T.R., Hansen E.K.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
DR EMBL: AE007510; AAK76279.1;
DR TIGR: SP2231;
KW Complete proteome.
SQ SEQUENCE 850 AA; 97303 MW; 1ADED613F06B5115 CRC64;

Query Match 30.4%; Score 69; DB 16; Length 850;
Best Local Similarity 33.3%; Pred. No. 2.5;
Matches 19; Conservative 10; Mismatches 16; Indels 12; Gaps 2;

QY 2 KKTLSKNDKFEIKTDELEIIIGSGSLSTFFRLNRSFTQALGK 46
Db 692 KKTFTDNVFSFENLGYTKKTFKNVNSFPGNSQVSESTFFRLDKTFTEAIOK 748

RESULT 5
Q9L651
ID Q9L651 PRELIMINARY; PRT; 648 AA.
AC Q9L651;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL LACTOCOCCIN 972 IMMUNITY PROTEIN (LACTIS PLASMID PBL1 DNA
DE FOR LACTOCOCCIN 972 OPERON).
GN LCLB.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Plasmid pBL1.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IPLA 972; PLASMID=PBL1;
RA Sanchez C., Hernandez de Rojas A., Martinez B., Arguelles M.E.,
RA Suarez J.E., Rodriguez A., Mayo B.;
RT "Nucleotide sequence of pBL1, a bacteriocin-producing plasmid from
RT Lactococcus lactis IPLA 972.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=IPLA 972; PLASMID=PBL1;
RA Suarez J.E.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=IPLA 972; PLASMID=PBL1;
RX MEDLINE=20055640; PubMed=10589723;
RA Martinez B., Fernandez M., Rodriguez A., Suarez J.E.;
RT "Synthesis of Lactococcin 972, a bacteriocin produced by Lactococcus
RT lactis IPLA 972, depends on the expression of a plasmid-encoded

```



```

RT biocistronic operon."
RL Microbiology 145:3155-3161(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=IPLA 972; PLASMID=PBL1;
RA Martinez B.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF242367; AAF64055.1; -
DR EMBL; AF002203; AAC03468.1; -
KW Plasmid.
SQ SEQUENCE 648 AA; 74139 MW; 79530E501D6371F7 CRC64;

Query Match 28.08; Score 63.5; DB 2; Length 648;
Best Local Similarity 46.28; Pred. No. 9.6;
Matches 18; Conservative 8; Mismatches 10; Indels 3; Gaps 2;

QY 5 SLKNDKFEIKTDELEIIIGSGSL--STF-FRLFNRSF 40
Db 528 LSVKNRFEIKSQQLFIWTTSSIALISSTLFVLMNKIY 566

RESULT 6
Q9CPE8 PRELIMINARY; PRT; 275 AA.
AC Q9CPE8;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROC OR PM0095.
GN Pasteurella multocida.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70."
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006044; AA02179.1; -
DR InterPro; IPR000304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR PROSITE; PS00521; P5CR; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 275 AA; 29556 MW; 011C7F1AD31A7D7 CRC64;

Query Match 26.28; Score 59.5; DB 16; Length 275;
Best Local Similarity 29.68; Pred. No. 12;
Matches 16; Conservative 7; Mismatches 16; Indels 15; Gaps 1;

QY 6 SLKNDKFEIKTD-----ELIIIGSGSLSTFFRLFNRSFTQA 44
Db 141 SLKPEYKQFTQDLLNAVNGKTCWQTQEAADMTITAGSGSPAYFLFMEAMQQA 194

RESULT 7
Q33596 PRELIMINARY; PRT; 48 AA.
AC Q33596;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PEPTIDE PRECURSOR.
GN COMC.
OS Streptococcus anginosus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1328;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 10713;

```

```

RX MEDLINE=98012953; PubMed=9352904;
RA Havarstein L.S., Hakenbeck R., Gaustad P.;
RT "Natural competence in the genus Streptococcus: Evidence that
streptococci can change phenotype by interspecies recombinational
exchanges."
RL J. Bacteriol. 179:6589-6594(1997).
DR EMBL; AJ000864; CAA04341.1; -
DR InterPro; IPR004288; COMC.
DR Pfam; PF03047; COMC; 1.
KW Signal.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 48 COMPETENCE STIMULATING PEPTIDE.
SQ SEQUENCE 48 AA; 5621 MW; E86DF7AEDE4E94DB CRC64;

Query Match 25.68; Score 58; DB 2; Length 48;
Best Local Similarity 30.28; Pred. No. 2.8;
Matches 16; Conservative 9; Mismatches 16; Indels 12; Gaps 2;

QY 1 MKKTLSLKN-----DFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQALCK 46
Db 1 MKKLFAPKKEVVKVEFEKELNDEQLDKIIIGSDSRIMGF-----DFSLFKK 48

RESULT 8
Q99WU3 PRELIMINARY; PRT; 1009 AA.
AC Q99WU3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SA0272 PROTEIN (HYPOTHETICAL PROTEIN SAV0283).
GN SA0272 OR SAV0283.
OS Staphylococcus aureus (strain N315), and
Staphylococcus aureus (strain Mu50).
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=158879, 158878;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain N315), and S.aureus (strain Mu50);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizutani-U I., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003130; BAB41496.1; -
DR EMBL; AP003358; BAB56445.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1009 AA; 114781 MW; 16AD06D170CA31FE CRC64;

Query Match 25.68; Score 58; DB 16; Length 1009;
Best Local Similarity 34.38; Pred. No. 79;
Matches 12; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 9 NDFKEIKTDELEIIIGSGSLSTFFRLFNRSFTQA 43
Db 202 NDFPELFTDVLNVSISANKDITKWFQTYNKSLLSA 236

RESULT 9
Q9CIA3 PRELIMINARY; PRT; 453 AA.
AC Q9CIA3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE SULFATE TRANSPORTER SUL2-LA (FRAGMENT).

```

```

OS Saccharomyces pastorianus (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=27292;
RN [1]
RP SEQUENCE FROM N.A.
RA James A.B., Slaughter C., Meaden P.G.;
RT "Characterization of partial coding sulfate transporter sequences from
RT Saccharomyces pastorianus and Saccharomycetes bayanus."
RL SUBMITTED (MAR-2001) TO THE EMBL/GenBank/DBJ databases.
DR EMBL; AF364410; AAK32879.1;
DR InterPro: IPR001902; Sulfate_transp.
DR Pfam: PF00916; Sulfate_transp; 1.
FT NON_TER 1
FT NON_TER 453
SQ SEQUENCE 453 AA; 49519 MW; 137117A0D12BD31F CRC64;

Query Match 25.3%; Score 57.5; DB 3; Length 453;
Best Local Similarity 36.7%; Pred. No. 38;
Matches 18; Conservative 6; Mismatches 18; Indels 7; Gaps 2;

QY 1 MKKTLKNDKFKIKDELEILLIGSGSLSTFFRL-----FNRSTQA 43
Db 227 ISKSGRVNDYKVPQDEL-IAIGVSNLLGTFFNAYPATGFSRSALKA 274

RESULT 10
ID Q9SUK0 PRELIMINARY; PRT; 516 AA.
AC Q9SUK0;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE GLUCOSE-6-PHOSPHATE 1-DEHYDROGENASE (EC 1.1.1.49) (G6PD).
GN ACG9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Wentz U.K., Hauschild R., Lange C., Pietersma M., Wenderoth I.,
RA von Schaeven A.;
RT "Evidence for functional convergence of redox regulation in G6PDH
RT Isoforms of cyanobacteria and higher plants."
RL Plant Mol. Biol. 40:487-494(1999).
CC -|- CATALYTIC ACTIVITY: D-GLUCOSE 6-PHOSPHATE + NADP(+) = D-GLUCONO-
CC DELTA-LACTONE 6-PHOSPHATE + NADPH.
CC -|- PATHWAY: FIRST STEP IN PENTOSE PHOSPHATE PATHWAY.
CC -|- SIMILARITY: BELONGS TO THE GLUCOSE-6-PHOSPHATE DEHYDROGENASE
CC FAMILY.
DR EMBL; AJ010970; CAB52674.1; -.
DR HSSP; P11411; LDPC.
DR InterPro: IPR001282; G6PD.
DR Pfam: PF00479; G6PD; 1.
DR Pfam: PF02781; G6PD.C; 1.
DR PRINTS; PR00079; G6PDHGRNASE.
DR ProDom; PD001129; G6PD; 1.
DR PROSITE; PS00069; G6P_DEHYDROGENASE; 1.
KW Glucose metabolism; NADP; Oxidoreductase.
SQ SEQUENCE 516 AA; 59171 MW; A7623780F053CA1C CRC64;

Query Match 25.3%; Score 57.5; DB 10; Length 516;
Best Local Similarity 41.5%; Pred. No. 44;
Matches 22; Conservative 7; Mismatches 11; Indels 13; Gaps 5;

QY 1 MKKTLKND--FKE-----IKTDELEIII-GSGSLS---TF---FRLFNRSF 40
Db 8 MEKRSLKNDSEVKEYNPVTETGSLSIIVLGASGDLAKKTFPALNLFHQGF 60

RESULT 12
ID Q9AHY9 PRELIMINARY; PRT; 591 AA.
AC Q9AHY9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE YFAA (FRAGMENT).
GN YFAA.
OS Photobacterium luminescens (Xenorhabdus luminescens).
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Photobacterium.
OX NCBI_TaxID=29488;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NC19, ATCC29304;
RL MEDLINE=21225535; PubMed=11325940;
RA Cliche T.A., Bintrim S.B., Horswill A.R., Ensign J.C.;
RT "A Phosphatethiethyl Transferrase Homolog Is Essential for
RT Photobacterium luminescens to Support Growth and Reproduction of the
RT Entomopathogenic Nematode Heterorhabditis bacteriophora."

```

```

RL J. Bacteriol. 183:3117-3126(2001).
DR EMBL; AF288085; AAK16098.1; -.
DR HSP; P39435; IKAS.
DR InterPro; IPR000794; Ketoacyl-synt.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR PFam; PF02801; ketoacyl-synt_C; 1.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
FT NON_TER 591
SQ SEQUENCE 591 AA; 63878 MW; AF1451C6DD500891 CRC64;

Query Match 25.1%; Score 57; DB 2; Length 591;
Best Local Similarity 34.3%; Pred. No. 59;
Matches 12; Conservative 7; Mismatches 16; Indels 0; Gaps 0;

QY 12 KEIKTDELEIIIGSGSLSTFFRLFNRSFTQALGK 46
   : : : : : : : : : : : : : : : : : :
Db 149 QRVNRDRIGVILGGNGNTLFLSLAGRQQTPLYLRK 183

RESULT 13
Q9RQ06 PRELIMINARY; PRT; 407 AA.
AC Q9RQ06;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BUSAA.
GN BUSAA.
OS Lactococcus lactis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1358;
RN [1]
RP SEQUENCE FROM N.A.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_Gtp_A.
DR InterPro; IPR000644; CBS.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport.
SQ SEQUENCE 407 AA; 45653 MW; 656D9097AE58407F CRC64;

Query Match 24.9%; Score 56.5; DB 2; Length 407;
Best Local Similarity 45.7%; Pred. No. 46;
Matches 16; Conservative 3; Mismatches 15; Indels 1; Gaps 1;

QY 13 EIKTDELEIIIGSGS-LSTFFRLFNRSFTQALGK 46
   || : : : : : ||| ||| ||| |||
Db 50 EINEGEIVFMGLSGSGKSTLLRLNRLIEPTSGK 84

RESULT 15
Q9KIF7 PRELIMINARY; PRT; 408 AA.
ID Q9KIF7
AC Q9KIF7;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE OPURA (BETAINE ABC TRANSPORTER ATP BINDING PROTEIN).
DE OPURA OR BUSAA.
GN Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20318987; PubMed=10860977;
RA van der Heide T., Poolman B.;
RT "Osmoregulated ABC-transport system of Lactococcus lactis senses water
   stress via changes in the physical state of the membrane.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:7102-7106(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21235186; PubMed=11337471;
RA Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
   Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
   lactis ssp. lactis IL1403.";
RL Genome Res. 11:731-753(2001).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
   (ABC TRANSPORTERS).
DR EMBL; AF234619; AAF37878.1; -.
DR EMBL; AE006375; AAK05550.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001687; ATP_Gtp_A.
DR Pfam; PF00005; ABC_tran; 1.
DR Pfam; PF00571; ABC; 2.

```

DR SMART; SM00382; AAA; 1.
DR SMART; SM00116; CBS; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Complete_proteome; Transport.
SQ SEQUENCE 408 AA; 45750 MW; EBB8A887BE563B88 CRC64;

Query Match 24.9%; Score 56.5; DB 16; Length 408;
Best Local Similarity 45.7%; Pred. No. 46;
Matches 16; Conservative 3; Mismatches 15; Indels 1; Gaps 1;

QY 13 EIKTDELEIIIGSGS-LSTFFRLFNRSFTQALGK 46
|| | : | | | | | | | | | |
Db 50 EINEGEIFVIMGLSGCKSTLLRLNRLIEPTSGK 84

Search completed: November 5, 2002, 10:57:30
Job time : 34.5224 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:57:42 ; Search time 36.3881 Seconds
(without alignments)
140.414 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 46

Sequence: 1 MKKTLKNDKFKETDELE.....GSLSTFFRLFNRSQTALGK 46

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 1000 summaries

Database : A_Geneseq_032802.*

1: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	15.2	115	22	AAO00376 Human polypeptide
2	7	15.2	118	22	AAO01734 Human polypeptide
3	7	15.2	126	22	AAO00422 Human polypeptide
4	7	15.2	150	22	AAO11182 Human polypeptide
5	7	15.2	186	21	AAI19904 Arabidopsis thalia
6	7	15.2	257	21	AAG19903 Arabidopsis thalia
7	7	15.2	287	21	AAG19902 Arabidopsis thalia
8	7	15.2	558	14	AAR42086 Mouse RelB protein
9	7	15.2	586	21	AAG30638 Arabidopsis thalia
10	7	15.2	612	21	AAG30637 Arabidopsis thalia
11	7	15.2	618	21	AAG30636 Arabidopsis thalia

12	7	15.2	658	22	ABG17981	Novel human diagno
13	7	15.2	767	22	ABB58240	Drosophila melanog
14	7	15.2	929	22	ABB70857	Drosophila melanog
15	7	15.2	2771	22	ABB63880	Drosophila melanog
16	7	15.2	17	16	AAR67567	Antagonist (18) of
17	6	13.0	19	21	AAI85614	Cathepsin-D antige
18	6	13.0	20	18	AAW33042	Human heat shock p
19	6	13.0	20	18	AAW12350	Human hsp60 peptid
20	6	13.0	27	20	AAI40080	Peptide sequence d
21	6	13.0	30	22	ABB29081	Peptide #1732 enco
22	6	13.0	30	22	ABB34240	Peptide #1746 enco
23	6	13.0	30	22	ABB19676	Protein #1675 enco
24	6	13.0	30	22	AAI55035	Human brain expres
25	6	13.0	30	22	AAW67423	Human bone marrow
26	6	13.0	30	22	AAI15251	Peptide #1685 enco
27	6	13.0	30	22	AAW27112	Peptide #1749 enco
28	6	13.0	30	22	AAW02995	Peptide #1677 enco
29	6	13.0	34	22	AAW91358	Human immune/haema
30	6	13.0	48	22	AAW23675	Human EST encoded
31	6	13.0	51	17	AAW94979	Calcium channel is
32	6	13.0	51	17	AAW94978	Calcium channel is
33	6	13.0	54	22	AAW77003	Human colon cancer
34	6	13.0	61	22	AAU46669	Propionibacterium
35	6	13.0	62	17	AAW97851	Rat brain calcium
36	6	13.0	66	17	AAW97852	Rat brain calcium
37	6	13.0	66	20	AAW60064	Human endometrium
38	6	13.0	66	22	ABG04703	Novel human diagno
39	6	13.0	69	22	ABB32283	Peptide #4934 enco
40	6	13.0	69	22	ABB22836	Protein #4835 enco
41	6	13.0	69	22	AAW58198	Human brain expres
42	6	13.0	69	22	AAW70656	Human bone marrow
43	6	13.0	69	22	AAW30969	Peptide #5006 enco
44	6	13.0	70	22	AAW92947	Human digestive sy
45	6	13.0	70	22	AAU20043	Human liver associ
46	6	13.0	72	22	AAO13104	Human polypeptide
47	6	13.0	78	22	AAW89377	Human immune/haema
48	6	13.0	82	22	ABB16579	Human nervous syst
49	6	13.0	96	19	AAW59796	Nucleotide sequenc
50	6	13.0	96	22	AAU58723	Propionibacterium
51	6	13.0	97	22	AAU45726	Propionibacterium
52	6	13.0	97	22	AAU46334	Propionibacterium
53	6	13.0	101	21	AAW57131	Human prostate can
54	6	13.0	105	22	ABG13791	Novel human diagno
55	6	13.0	109	22	AAO09578	Human polypeptide
56	6	13.0	110	20	AAI12436	Human polypeptide
57	6	13.0	111	22	AAO10094	Human polypeptide
58	6	13.0	112	22	AAU43924	Human polypeptide
59	6	13.0	118	22	AAO01659	Synthetic autoanti
60	6	13.0	121	22	AAU02075	Human polypeptide
61	6	13.0	121	22	AAO06498	Peptide sequence d
62	6	13.0	122	20	AAU40074	Peptide sequence d
63	6	13.0	126	21	AAW08937	H. pylori cytoplas
64	6	13.0	132	18	AAW20491	Human secreted pro
65	6	13.0	132	22	AAU20354	E. canis lipoprote
66	6	13.0	139	22	AAW62816	Amino acid sequenc
67	6	13.0	141	20	AAW35818	Propionibacterium
68	6	13.0	141	22	AAU45360	Human CCAAT/enhanc
69	6	13.0	147	22	AAE11946	Novel human diagno
70	6	13.0	149	22	ABG17274	Human novel secret
71	6	13.0	150	22	AAU16099	Arabidopsis thalia
72	6	13.0	152	21	AAW48562	Arabidopsis thalia
73	6	13.0	159	21	AAW08936	Human novel secret
74	6	13.0	172	22	AAU16362	Human polypeptide
75	6	13.0	177	22	AAO01381	Amino acid sequenc
76	6	13.0	180	19	AAW59797	Human diagnostic a
77	6	13.0	194	22	AAU19473	Arabidopsis thalia
78	6	13.0	198	21	AAW03322	Arabidopsis thalia
79	6	13.0	198	21	AAW52804	Rat proteosome del
80	6	13.0	202	15	AAW47470	Novel human diagno
81	6	13.0	212	22	ABG24830	RING12 proteasome
82	6	13.0	219	13	AAW25592	Rat proteasome RIN
83	6	13.0	219	15	AAW47473	Human proteasome s
84	6	13.0	219	17	AAW06438	

85	6	13.0	220	19	AAW98473	H. pylori GHPO 840	158	6	13.0	360	21	AAW69364	Amino acid sequenc
86	6	13.0	222	20	AAW37757	Chlamydia trachoma	159	6	13.0	360	22	ABW70816	Drosophila melanog
87	6	13.0	223	21	AAW09321	Arabidopsis thalia	160	6	13.0	360	22	ABG14697	Novel human diago
88	6	13.0	224	21	AAW52803	Arabidopsis thalia	161	6	13.0	360	22	AAW41558	Human polypeptide
89	6	13.0	224	22	ABW61835	Arabidopsis thalia	162	6	13.0	360	22	AAW82528	S. epidermidis ope
90	6	13.0	233	21	AAW09320	Arabidopsis thalia	163	6	13.0	365	22	AAW25511	Human protein sequ
91	6	13.0	233	21	AAW52802	Arabidopsis thalia	164	6	13.0	368	22	AAW01696	Human gene 1 encod
92	6	13.0	239	16	AAW83970	Human proteasome s	165	6	13.0	371	22	AAW40429	Human polypeptide
93	6	13.0	239	17	AAW06435	Human proteasome s	166	6	13.0	372	22	AAW48589	Human polypeptide
94	6	13.0	243	20	AAW34866	Chlamydia pneumoni	167	6	13.0	372	22	AAW52301	Protonibacterium
95	6	13.0	244	22	AAW81950	S. epidermidis ope	168	6	13.0	376	22	AAW82506	ActA protein fragm
96	6	13.0	247	22	AAW63519	Human gastric can	169	6	13.0	376	22	AAW63692	S. epidermidis ope
97	6	13.0	252	22	ABW43517	Novel human diago	170	6	13.0	377	22	AAW63692	Human secreted pro
98	6	13.0	252	22	ABW43517	Novel human diago	171	6	13.0	377	22	AAW93203	Human polypeptide
99	6	13.0	254	22	ABW12787	Novel human diago	172	6	13.0	377	22	AAW38833	Human polypeptide
100	6	13.0	254	22	ABW12787	Novel human diago	173	6	13.0	377	22	AAW01242	Human protein diff
101	6	13.0	260	22	AAW96770	Putative P. abyssi	174	6	13.0	379	13	AAW29630	Apase protein. Fu
102	6	13.0	269	15	AAW46337	PB92 serine protea	175	6	13.0	382	21	AAW45967	Arabidopsis thalia
103	6	13.0	269	15	AAW46338	Subtilisin 309 Ser	176	6	13.0	384	22	AAW24101	Human EST encoded
104	6	13.0	269	15	AAW46339	PB92 serine protea	177	6	13.0	385	21	AAW45966	Arabidopsis thalia
105	6	13.0	269	15	AAW46340	Subtilisin 309 Ser	178	6	13.0	385	22	AAW40619	Human polypeptide
106	6	13.0	269	15	AAW46341	PB92 serine protea	179	6	13.0	387	21	AAW45965	Arabidopsis thalia
107	6	13.0	269	15	AAW46342	Subtilisin 309 Ser	180	6	13.0	392	22	ABW58237	Drosophila melanog
108	6	13.0	269	15	AAW46343	PB92 serine protea	181	6	13.0	396	22	AAW04799	Neurospora crassa
109	6	13.0	269	15	AAW46344	Subtilisin 309 Ser	182	6	13.0	397	18	AAW31628	Aspergillus oryzae
110	6	13.0	270	13	AAW27641	Human calcium chan	183	6	13.0	397	22	AAW04886	Human protease pro
111	6	13.0	273	18	AAW20926	H. pylori cytoplas	184	6	13.0	398	16	AAW75299	Aspergillus niger
112	6	13.0	276	22	ABW69618	Drosophila melanog	185	6	13.0	398	22	ABG10841	Novel human diago
113	6	13.0	277	22	AAW25545	Human protein sequ	186	6	13.0	404	22	ABW64873	Drosophila melanog
114	6	13.0	279	22	ABW16367	Human nervous syst	187	6	13.0	404	22	AAW75543	Human colon cancer
115	6	13.0	279	22	AAW16524	Human novel secret	188	6	13.0	410	22	AAW01672	Human gene 1 encod
116	6	13.0	282	22	AAW05262	Chicken stem cell	189	6	13.0	412	16	AAW74207	Human death associ
117	6	13.0	282	22	AAW98364	Chicken SCF protei	190	6	13.0	412	19	AAW71369	Death associated p
118	6	13.0	282	22	AAW02489	Chicken SCF (stem	191	6	13.0	412	20	AAW06478	Human tumour-assoc
119	6	13.0	282	22	AAW02775	Chicken SCF (stem	192	6	13.0	412	21	AAW48540	Helicobacter pylor
120	6	13.0	282	22	AAW73576	Chicken SCF (stem	193	6	13.0	412	21	AAW93685	Mature secreted hum
121	6	13.0	282	22	AAW96949	Chicken stem cell	194	6	13.0	414	21	AAW48539	Arabidopsis thalia
122	6	13.0	283	18	AAW20637	H. pylori flagella	195	6	13.0	423	22	ABW62432	Arabidopsis thalia
123	6	13.0	286	22	AAW32713	Novel human secret	196	6	13.0	439	21	AAW19115	Drosophila melanog
124	6	13.0	289	22	ABW60987	Drosophila melanog	197	6	13.0	442	19	AAW98707	Polypeptide isolat
125	6	13.0	289	22	ABW23002	Novel human diago	198	6	13.0	442	22	AAW35813	H. pylori GHPO 728
126	6	13.0	296	22	AAE11947	Mouse CCAAT/enhanc	199	6	13.0	446	20	AAW30792	Helicobacter pylor
127	6	13.0	297	22	AAE11948	Rat CCAAT/enhancer	200	6	13.0	447	21	AAW29728	Mature secreted hum
128	6	13.0	306	22	AAW96808	Putative P. abyssi	201	6	13.0	447	21	AAW48538	Arabidopsis thalia
129	6	13.0	309	22	AAW96496	Putative P. abyssi	202	6	13.0	449	21	AAW30230	Arabidopsis thalia
130	6	13.0	312	22	ABW08404	Novel human diago	203	6	13.0	452	20	AAW00864	Arabidopsis thalia
131	6	13.0	317	22	ABW69511	Drosophila melanog	204	6	13.0	452	20	AAW95029	Cellulohydrolyase
132	6	13.0	321	22	AAU34275	Staphylococcus aur	205	6	13.0	452	21	AAW14049	A. niger cellolohio
133	6	13.0	321	22	AAU37101	Staphylococcus aur	206	6	13.0	452	21	AAW51985	Arabidopsis thalia
134	6	13.0	321	22	AAU37128	Staphylococcus aur	207	6	13.0	455	21	AAW17438	Arabidopsis thalia
135	6	13.0	322	21	AAW19116	Homologue of polyp	208	6	13.0	455	21	AAW21851	Arabidopsis thalia
136	6	13.0	322	22	AAE11945	Human CCAAT/enhanc	209	6	13.0	455	22	AAW90723	C glutamicum prote
137	6	13.0	329	22	AAW35549	Haemophilus influe	210	6	13.0	456	20	AAW17067	Human 3-OST-4 prot
138	6	13.0	329	22	ABG11238	Novel human diago	211	6	13.0	459	22	AAU02077	Synthetic multi-ta
139	6	13.0	330	22	AAW96503	Putative P. abyssi	212	6	13.0	463	18	AAW55714	H. pylori ORF 09cp
140	6	13.0	331	22	AAW95781	5H7 single chain a	213	6	13.0	465	22	ABG03158	Novel human diago
141	6	13.0	331	22	ABG19411	Novel human diago	214	6	13.0	466	20	AAW30793	Immature human HKN
142	6	13.0	333	22	AAW72693	Murine OR-like pol	215	6	13.0	466	21	AAW17437	Arabidopsis thalia
143	6	13.0	334	22	AAW67978	Protonibacterium	216	6	13.0	466	21	AAW21850	Arabidopsis thalia
144	6	13.0	339	22	ABW26859	Novel human diago	217	6	13.0	468	21	AAW19042	Amino acid sequenc
145	6	13.0	342	22	ABW68951	Drosophila melanog	218	6	13.0	468	21	AAW29727	Arabidopsis thalia
146	6	13.0	342	22	AAU09816	Fibroblast growth	219	6	13.0	472	22	AAU09810	Human fibroblast g
147	6	13.0	344	22	ABG03199	Novel human diago	220	6	13.0	473	21	AAW51984	Arabidopsis thalia
148	6	13.0	344	22	AAW81114	Mycobacterium tube	221	6	13.0	474	21	AAW14048	Arabidopsis thalia
149	6	13.0	345	12	AAW14408	Nuclear factor C/E	222	6	13.0	476	13	AAW20540	Alkaline protease.
150	6	13.0	345	22	AAE11944	Human CCAAT/enhanc	223	6	13.0	477	20	AAW30786	Protein encoded by
151	6	13.0	345	22	AAU09068	Human transcriptio	224	6	13.0	477	20	AAW97105	Thermostable facto
152	6	13.0	350	22	AAW52302	ActA protein fragm	225	6	13.0	477	22	AAW94726	Human protein sequ
153	6	13.0	350	22	AAW85789	Human kinase PKIN-	226	6	13.0	478	22	ABW58996	Drosophila melanog
154	6	13.0	351	18	AAW20341	H. pylori flagella	227	6	13.0	481	22	AAW96058	Putative P. abyssi
155	6	13.0	355	18	AAW20737	H. pylori flagella	228	6	13.0	489	13	AAW27647	Human calcium chan
156	6	13.0	356	22	AAW39772	Human polypeptide	229	6	13.0	489	21	AAW17436	Arabidopsis thalia
157	6	13.0	359	22	ABW70995	Drosophila melanog	230	6	13.0	493	22	AAE07168	Mycobacterium tube
	6	13.0	360	20	AAW34924	Chlamydia pneumoni		6	13.0	495	20	AAW30785	Protein encoded by

231	6	13.0	496	20	AAY37175	Amino acid sequenc	304	6	13.0	719	22	AAB74219	Bacillus maltogeni
232	6	13.0	496	22	AAG80007	D. melanogaster li	305	6	13.0	719	22	AAB74220	Bacillus maltogeni
233	6	13.0	497	21	AAG21849	Arabidopsis thalia	306	6	13.0	719	22	AAB74221	Bacillus maltogeni
234	6	13.0	499	13	AAR27574	The ABP-B from A.	307	6	13.0	719	22	AAB74222	Bacillus maltogeni
235	6	13.0	501	22	ABB64937	Drosophila melanog	308	6	13.0	719	22	AAB74223	Bacillus maltogeni
236	6	13.0	504	21	AAB24066	Human PRO943 prote	309	6	13.0	755	22	AAR23150	Mutant thermostabl
237	6	13.0	504	21	AAY92864	Human fibroblast g	310	6	13.0	766	22	AAY34979	Enterococcus faeca
238	6	13.0	504	21	AAY66656	Membrane-bound pro	311	6	13.0	785	22	AAG91691	C glutamicum prote
239	6	13.0	504	22	AAB65179	Human PRO943 (UNQ4	312	6	13.0	796	22	ABG01448	Novel human diagno
240	6	13.0	504	22	AAB65264	Human polypeptide	313	6	13.0	809	22	ABG11797	Novel human diagno
241	6	13.0	518	22	AAM41685	Human calcium chan	314	6	13.0	821	13	ABG23149	Drosophila melanog
242	6	13.0	518	22	AAR27644	Drosophila melanog	315	6	13.0	821	13	AAR23149	Mutant thermostabl
243	6	13.0	526	22	ABBS8656	Polypeptide isolat	316	6	13.0	821	13	ABG21079	Novel human diagno
244	6	13.0	529	21	ABJ19114	Murine fibroblast	317	6	13.0	844	22	ABG63141	Novel human diagno
245	6	13.0	529	22	AAU09808	Expressed antigen	318	6	13.0	857	13	AAG45432	Drosophila melanog
246	6	13.0	530	20	AAW90018	Human MLN 51. Hom	319	6	13.0	860	21	AAR23147	Mutant thermostabl
247	6	13.0	534	18	AAW25769	Human chaperone pr	320	6	13.0	862	22	ABG23271	Balanus amphitrite
248	6	13.0	538	18	AAW10026	Partial sporozoite	321	6	13.0	875	22	ABBS9530	Drosophila melanog
249	6	13.0	547	16	AAR67385	Mitochondrial prot	322	6	13.0	882	13	AAR23148	Mutant thermostabl
250	6	13.0	548	20	AAU06927	C. albicans antige	323	6	13.0	887	15	AAR58609	Novel human diagno
251	6	13.0	553	22	ABG61472	Drosophila melanog	324	6	13.0	888	20	AAY49560	Hamster HMG-CoA re
252	6	13.0	555	13	AAR22925	Truncated hamster	325	6	13.0	893	13	AAR21830	Human protein C pr
253	6	13.0	559	21	AAV91945	Human chaperone pr	326	6	13.0	893	13	AAR23146	Sequence of a ther
254	6	13.0	559	22	AAW93115	Human protein sequ	327	6	13.0	893	20	AAW78469	Mutant thermostabl
255	6	13.0	560	19	AAW98592	H. pylori GPO 97	328	6	13.0	908	22	AAW78469	Thermatoga maritim
256	6	13.0	561	22	AAU34618	E. coli cellular p	329	6	13.0	913	22	AAW00834	Human bone marrow
257	6	13.0	567	21	AAG30229	Arabidopsis thalia	330	6	13.0	913	22	AAW53829	Pseudomonas Orfx s
258	6	13.0	571	22	AAU36220	Pseudomonas aerugi	331	6	13.0	913	21	AAW82595	Pseudomonas alcali
259	6	13.0	573	11	AAR04713	Amino acid sequenc	332	6	13.0	913	22	AAW82595	Pseudomonas alcali
260	6	13.0	573	18	AAW14946	Human heat shock p	333	6	13.0	917	22	ABG06296	Novel human diagno
261	6	13.0	573	18	AAW12345	Human heat shock p	334	6	13.0	917	22	ABG06296	Novel human diagno
262	6	13.0	573	18	AAW01657	Human heat shock p	335	6	13.0	926	22	ABG21080	Novel human diagno
263	6	13.0	573	19	AAW56120	Protein sequence o	336	6	13.0	951	22	AAW79950	Human protein SEQ
264	6	13.0	573	20	AAV23926	Amino acid sequenc	337	6	13.0	955	21	AAW42231	Human ORFX ORF1995
265	6	13.0	573	21	AAV93333	Amino acid sequenc	338	6	13.0	960	21	AAG45431	Arabidopsis thalia
266	6	13.0	573	21	AAV93333	Amino acid sequenc	339	6	13.0	962	22	ABG11976	Human TRAP150 homo
267	6	13.0	573	22	ABBS0262	Heat shock protein	340	6	13.0	977	22	AAW51671	Phaseolus lunatus
268	6	13.0	573	22	AAE11756	Human heat shock p	341	6	13.0	986	21	AAG45430	Arabidopsis thalia
269	6	13.0	575	16	AAR64763	Human PI protein.	342	6	13.0	993	13	AAR27650	Human calcium chan
270	6	13.0	580	22	ABBS5504	Drosophila melanog	343	6	13.0	1015	22	ABG08785	Novel human diagno
271	6	13.0	588	22	ABG12546	Novel human diagno	344	6	13.0	1053	22	AAW78966	Human protein SEQ
272	6	13.0	594	22	AAU09817	Murine FGFR-L extr	345	6	13.0	1086	22	ABG60786	Drosophila melanog
273	6	13.0	595	22	AAU38334	Salmonella typhi c	346	6	13.0	1127	22	ABG24996	Novel human diagno
274	6	13.0	595	22	ABB32176	Peptide #4827 enco	347	6	13.0	1174	22	ABG1704	Novel human diagno
275	6	13.0	595	22	ABB22720	Protein #4719 enco	348	6	13.0	1217	22	AAU33177	Drosophila melanog
276	6	13.0	597	21	AAV90257	Streptococcus equi	349	6	13.0	1218	22	ABG63445	Novel human secret
277	6	13.0	600	21	AAW40930	Zea mays protein f	350	6	13.0	1222	22	ABG66096	Drosophila melanog
278	6	13.0	608	22	AAW38643	Human polypeptide	351	6	13.0	1237	22	AAW20014	Drosophila melanog
279	6	13.0	610	22	AAW52300	ActA protein. Lis	352	6	13.0	1342	22	ABG4632	Drosophila SREBP c
280	6	13.0	611	13	AAR23151	Mutant thermostabl	353	6	13.0	1343	22	AAW00947	Drosophila melanog
281	6	13.0	611	21	AAB32732	Eucalyptus grandis	354	6	13.0	1359	21	AAW41785	Human bone marrow
282	6	13.0	611	22	ABB71064	Drosophila melanog	355	6	13.0	1371	22	ABBS2745	Human ORFX ORF1549
283	6	13.0	613	22	ABG27790	Novel human diagno	356	6	13.0	1437	22	AAW79507	Escherichia coli p
284	6	13.0	615	21	AAG30228	Arabidopsis thalia	357	6	13.0	1820	22	AAU27679	Human protein SEQ
285	6	13.0	635	19	AAW75855	Human secretory pr	358	6	13.0	1824	21	AAV49431	Human full-length
286	6	13.0	635	20	AAW45160	Human membrane tra	359	6	13.0	1856	20	AAV21801	Murine CACNAIF pro
287	6	13.0	635	21	AAW71060	Amino acid sequenc	360	6	13.0	1856	20	AAV21801	B. subtilis rib op
288	6	13.0	639	22	AAU09142	Drosophila melanog	361	6	13.0	1856	21	AAW83269	Polypeptide encode
289	6	13.0	639	22	AAW79171	Arabidopsis thalia	362	6	13.0	1897	22	AAU15088	Protein encoded by
290	6	13.0	643	22	ABG60087	Corn CCR4 transcri	363	6	13.0	1912	21	AAV49429	Human CACNAIF spli
291	6	13.0	645	21	AAU94657	Novel human diagno	364	6	13.0	1967	14	AAR33547	Sequence of the al
292	6	13.0	683	22	ABG06295	Human calcium chan	365	6	13.0	1968	19	AAW63139	Human calcium chan
293	6	13.0	684	13	AAR27645	Novel human diagno	366	6	13.0	1977	21	AAV49430	Human CACNAIF long
294	6	13.0	692	22	ABG08784	Human GRPase assoc	367	6	13.0	2006	22	ABW71277	Drosophila melanog
295	6	13.0	698	21	AAV99666	Sequence of lag D	368	6	13.0	2110	22	ABBS8077	Drosophila melanog
296	6	13.0	703	15	AAR49135	Amino acid sequenc	369	6	13.0	2138	16	AAR72607	Drosophila melanog
297	6	13.0	719	20	AAV30821	Bacillus sp. malto	370	6	13.0	2138	21	AAW10593	Human neuronal cal
298	6	13.0	719	20	AAV31731	Bacillus maltogeni	371	6	13.0	2161	14	AAR33545	Human calcium chan
299	6	13.0	719	21	AAV94271	Bacillus maltogeni	372	6	13.0	2161	16	AAR71001	Sequence of the al
300	6	13.0	719	22	AAW59956	Bacillus maltogeni	373	6	13.0	2161	16	AAR71002	Human neuronal cal
301	6	13.0	719	22	AAB74216	Bacillus maltogeni	374	6	13.0	2161	19	AAW63137	Human calcium chan
302	6	13.0	719	22	AAB74217	Bacillus maltogeni	375	6	13.0	2161	19	AAW63149	Human calcium chan
303	6	13.0	719	22	AAB74218	Bacillus maltogeni	376	6	13.0	2161	21	AAB10568	Human calcium chan

377	6	13.0	2163	16	AA871003	Human neuronal cal	450	5	10.9	10	22	AA897220	Human complementar
378	6	13.0	2163	21	AA810570	Human calcium chan	451	5	10.9	10	22	AA897222	Human complementar
379	6	13.0	2279	22	AA838733	Human polypeptide	452	5	10.9	10	22	AA897294	Human complementar
380	6	13.0	2297	22	AA840519	Human polypeptide	453	5	10.9	10	22	AA897991	Human complementar
381	6	13.0	2322	18	AA815566	Melanoma-associate	454	5	10.9	10	22	AA897992	Human complementar
382	6	13.0	2344	22	AA837120	Staphylococcus aur	455	5	10.9	10	22	AA897993	Human complementar
383	6	13.0	2778	22	AB858683	Drosophila melanog	456	5	10.9	10	22	AA897994	Human complementar
384	6	13.0	26926	22	AA805396	Human titin (conne	457	5	10.9	10	22	AA885591	Saccharomyces cere
385	5	10.9	5	14	AA844624	Caicitonin fragmen	458	5	10.9	10	22	AA885593	Proto-oncogene C-CB
386	5	10.9	7	19	AA870511	Escherichia coli g	459	5	10.9	10	22	AA893432	Proto-oncogene C-CB
387	5	10.9	8	14	AA873305	Human hepatocyte g	460	5	10.9	10	22	AA893433	Epstein-Barr virus
388	5	10.9	8	15	AA853960	N-terminal sequenc	461	5	10.9	11	7	AA860483	Synthetic random c
389	5	10.9	8	18	AA813431	Melanoma-homing pe	462	5	10.9	11	13	AA825060	HraIpha-I11-beta 1
390	5	10.9	8	19	AA860292	Melanoma tumour ho	463	5	10.9	11	22	AA864070	Amino acid sequenc
391	5	10.9	8	20	AA893716	Mouse B16B15b mela	464	5	10.9	12	22	AA863314	Sequence #4 from a
392	5	10.9	8	21	AA821704	Murine melanoma ho	465	5	10.9	13	14	AA840791	Mouse pancreatic-d
393	5	10.9	8	21	AA823539	Angiogenic vascula	466	5	10.9	14	20	AA806305	Human peptide #568
394	5	10.9	8	21	AA817961	Integrin-binding p	467	5	10.9	14	22	AA897293	Human peptide #991
395	5	10.9	8	22	AA806282	Tumour homing pept	468	5	10.9	14	22	AA897716	Multi-functional c
396	5	10.9	9	17	AA849247	Human leucocyte an	469	5	10.9	15	19	AA877802	Linker 2 used in hum
397	5	10.9	10	19	AA877801	Multi-functional c	470	5	10.9	15	19	AA869020	Linker 2 used for
398	5	10.9	10	19	AA869019	Gly/Ser-rich linke	471	5	10.9	15	20	AA846296	Fragment of dnab8 c
399	5	10.9	10	21	AA881276	Human complementar	472	5	10.9	15	20	AA849078	TPQ receptor bind1
400	5	10.9	10	22	AA894010	Human complementar	473	5	10.9	15	20	AA822379	Mouse beta-actin a
401	5	10.9	10	22	AA894012	Human complementar	474	5	10.9	15	20	AA892605	Beta-actin referen
402	5	10.9	10	22	AA894124	Human complementar	475	5	10.9	15	20	AA892534	Plasminogen activa
403	5	10.9	10	22	AA894128	Human complementar	476	5	10.9	16	15	AA847026	Human granulocytic
404	5	10.9	10	22	AA894261	Human complementar	477	5	10.9	16	20	AA848746	Rat ENDO-I protein fr
405	5	10.9	10	22	AA894265	Human complementar	478	5	10.9	16	20	AA881961	Rat HPR protein fr
406	5	10.9	10	22	AA894265	Human complementar	479	5	10.9	16	20	AA881962	DPVD containing pe
407	5	10.9	10	22	AA894277	Human complementar	480	5	10.9	17	21	AA845049	KIX domain binding
408	5	10.9	10	22	AA894277	Human complementar	481	5	10.9	17	22	AA860913	Linker peptide #7,
409	5	10.9	10	22	AA894281	Human complementar	482	5	10.9	18	22	AA871888	N-terminus of 37 k
410	5	10.9	10	22	AA894281	Human complementar	483	5	10.9	19	14	AA836604	Peptide fragment p
411	5	10.9	10	22	AA894301	Human complementar	484	5	10.9	19	19	AA869830	Bovine vascular sm
412	5	10.9	10	22	AA894303	Human complementar	485	5	10.9	19	20	AA804268	Human ENDO-I prote
413	5	10.9	10	22	AA894305	Human complementar	486	5	10.9	19	20	AA881955	Human ENDO-I varia
414	5	10.9	10	22	AA894307	Human complementar	487	5	10.9	19	20	AA881957	Epstein-Barr virus
415	5	10.9	10	22	AA894309	Human complementar	488	5	10.9	20	19	AA865563	Epstein-Barr virus
416	5	10.9	10	22	AA894311	Human complementar	489	5	10.9	20	22	AA809083	Epstein-Barr virus
417	5	10.9	10	22	AA894313	Human complementar	490	5	10.9	21	18	AA834053	Human MDRI-P glyco
418	5	10.9	10	22	AA894315	Human complementar	491	5	10.9	21	21	AA895779	Linker peptide. S
419	5	10.9	10	22	AA894317	Human complementar	492	5	10.9	21	22	ABB40487	Peptide #7993 enco
420	5	10.9	10	22	AA894323	Human complementar	493	5	10.9	21	22	ABB24820	Protein #6819 enco
421	5	10.9	10	22	AA894325	Human complementar	494	5	10.9	21	22	AA861312	Human brain expres
422	5	10.9	10	22	AA894333	Human complementar	495	5	10.9	21	22	AA861312	Human bone marrow
423	5	10.9	10	22	AA894390	Human complementar	496	5	10.9	21	22	AA820204	Peptide #6638 enco
424	5	10.9	10	22	AA894392	Human complementar	497	5	10.9	21	22	AA834211	Peptide #8248 enco
425	5	10.9	10	22	AA895196	Human complementar	498	5	10.9	21	22	AA897036	Chimeric polypepti
426	5	10.9	10	22	AA895198	Human complementar	499	5	10.9	21	22	ABB30607	Peptide #3258 enco
427	5	10.9	10	22	AA895256	Human complementar	500	5	10.9	22	22	ABB35771	Peptide #3277 enco
428	5	10.9	10	22	AA895316	Human complementar	501	5	10.9	22	22	ABB21195	Protein #3194 enco
429	5	10.9	10	22	AA895380	Human complementar	502	5	10.9	22	22	AA856580	Human brain expres
430	5	10.9	10	22	AA896214	Human complementar	503	5	10.9	22	22	AA868960	Human bone marrow
431	5	10.9	10	22	AA896214	Human complementar	504	5	10.9	22	22	AA829271	Peptide #3221 enco
432	5	10.9	10	22	AA896424	Human complementar	505	5	10.9	22	22	AA829271	Peptide #3308 enco
433	5	10.9	10	22	AA896426	Human complementar	506	5	10.9	22	22	AA804504	Peptide #3186 enco
434	5	10.9	10	22	AA896428	Human complementar	507	5	10.9	22	22	AA865099	Random biotinylati
435	5	10.9	10	22	AA896430	Human complementar	508	5	10.9	23	16	AA865099	Biotinylation pept
436	5	10.9	10	22	AA896444	Human complementar	509	5	10.9	23	19	AA846702	Biotinylation pept
437	5	10.9	10	22	AA896446	Human complementar	510	5	10.9	23	20	AA867910	Human secreted pro
438	5	10.9	10	22	AA896456	Human complementar	511	5	10.9	23	20	AA871142	Human secreted pro
439	5	10.9	10	22	AA896624	Human complementar	512	5	10.9	23	21	AA871142	Angiogenin inhibit
440	5	10.9	10	22	AA896626	Human complementar	513	5	10.9	23	22	AAE12642	ATP-binding casset
441	5	10.9	10	22	AA896628	Human complementar	514	5	10.9	23	22	AAU04311	Human gene 14 enco
442	5	10.9	10	22	AA897132	Human complementar	515	5	10.9	23	22	AAE06119	Biotinylated pepti
443	5	10.9	10	22	AA897136	Human complementar	516	5	10.9	23	22	AAE09937	Flexibile linker us
444	5	10.9	10	22	AA897176	Human complementar	517	5	10.9	24	21	AA893255	Human secreted pro
445	5	10.9	10	22	AA897188	Human complementar	518	5	10.9	24	22	AB850758	Novel human diagn
446	5	10.9	10	22	AA897206	Human complementar	519	5	10.9	25	22	ABG01293	Peptide #9374 enco
447	5	10.9	10	22	AA897210	Human complementar	520	5	10.9	25	22	ABB41868	Protein #7567 enco
448	5	10.9	10	22	AA897212	Human complementar	521	5	10.9	25	22	AA825568	Human brain expres
449	5	10.9	10	22	AA897216	Human complementar	522	5	10.9	25	22	AA862741	

523	5	10.9	25	22	AAW75560	Human bone marrow	596	42	22	ABG12103	Novel human diago
524	5	10.9	25	22	AAW20639	Peptide #7073 enco	597	42	22	ABG12668	Novel human diago
525	5	10.9	25	22	AAW35666	Peptide #9703 enco	598	42	22	AAB67213	HER2 peptide ligan
526	5	10.9	27	20	AAW67973	Fragment of human	599	43	20	AAY60379	Human normal blad
527	5	10.9	29	19	AAW79384	Staphylococcus aur	600	43	22	AAW86005	Human immune/haema
528	5	10.9	29	20	AAW38452	Human secreted pro	601	43	22	AAW82905	Human TNF receptor
529	5	10.9	29	22	ABW39191	Peptide #6697 enco	602	44	21	AAW27604	Human secreted pro
530	5	10.9	29	22	ABW24049	Protein #6048 enco	603	44	22	ABG12084	Novel human diago
531	5	10.9	29	22	AAW59850	Human brain expres	604	44	22	ABG28609	Novel human diago
532	5	10.9	29	22	AAW72439	Human bone marrow	605	44	22	AAO00262	Human polypeptide
533	5	10.9	29	22	AAW32688	Peptide #6725 enco	606	45	20	AAW88773	Polypeptide fragme
534	5	10.9	30	20	AAW42794	Streptococcus pyog	607	45	21	AAB63064	Human secreted pro
535	5	10.9	30	21	AAW09195	Hepatitis GB virus	608	45	22	ABW50606	Human secreted pro
536	5	10.9	30	22	AAW08982	Human polypeptide	609	45	22	AAW82748	Human immune/haema
537	5	10.9	31	7	AAW61348	Sequence of calcit	610	46	18	AAW44420	Nucleic acid deliv
538	5	10.9	31	7	AAW61349	Sequence of calcit	611	46	18	AAW40849	Human secreted pro
539	5	10.9	31	22	ABW29408	Peptide #2059 enco	612	46	22	AAW89492	Human immune/haema
540	5	10.9	31	22	ABW34590	Peptide #2096 enco	613	46	22	AAO09653	Human polypeptide
541	5	10.9	31	22	ABW19996	Protein #1995 enco	614	46	22	AAG76210	Human colon cancer
542	5	10.9	31	22	AAW55379	Human brain expres	615	47	22	ABW41431	Peptide #8937 enco
543	5	10.9	31	22	AAW57774	Human bone marrow	616	47	22	AAW62304	Human brain expres
544	5	10.9	31	22	AAW75470	Human bone marrow	617	47	22	AAW75107	Human bone marrow
545	5	10.9	31	22	AAW15583	Peptide #2017 enco	618	47	22	AAW35223	Peptide #9260 enco
546	5	10.9	31	22	AAW28075	Peptide #2112 enco	619	48	18	AAW27749	Amino acid sequenc
547	5	10.9	31	22	AAW03327	Peptide #2009 enco	620	48	22	AAO09043	Human polypeptide
548	5	10.9	32	17	AAW03618	Rat bradykinin GPR	621	48	22	AAO09243	Human polypeptide
549	5	10.9	32	22	AAW03546	Human polypeptide	622	48	22	AAO11779	Human polypeptide
550	5	10.9	32	22	AAW92939	C glutamic monome	623	48	22	AAG76454	Human colon cancer
551	5	10.9	34	18	AAW15081	Lac28E monomer use	624	49	16	AAW65749	Linker sequence us
552	5	10.9	34	21	AAW27608	Human secreted pro	625	49	19	AAW98307	H. pylori GPO 535
553	5	10.9	34	22	ABW08616	Novel human diago	626	49	21	AAW27809	Human secreted pro
554	5	10.9	35	15	AAW57184	GET epitope C inse	627	49	22	AAG37929	Arabidopsis thalia
555	5	10.9	35	21	AAW38405	Fragment of human	628	49	22	AAE14027	Chemically modifi
556	5	10.9	35	22	ABG15474	Novel human diago	629	49	22	ABW42096	Peptide #9602 enco
557	5	10.9	35	22	AAW18882	Peptide #5316 enco	630	49	22	ABW14832	Human nervous syst
558	5	10.9	36	15	AAW57187	GET epitope C inse	631	49	22	AAW22665	Novel human colon
559	5	10.9	36	21	ABW38977	Human secreted pep	632	49	22	AAW62976	Human brain expres
560	5	10.9	36	22	ABW10490	Novel human diago	633	49	22	AAW75789	Human bone marrow
561	5	10.9	37	22	ABW42942	Peptide #10448 enc	634	49	22	AAW82540	Human immune/haema
562	5	10.9	37	22	AAW63844	Human brain expres	635	49	22	AAW92671	Human digestive sy
563	5	10.9	37	22	AAW76657	Human bone marrow	636	49	22	AAO11940	Human polypeptide
564	5	10.9	37	22	AAW36762	Peptide #10799 enc	637	49	22	AAW35899	Peptide #9936 enco
565	5	10.9	38	15	AAW57185	GET epitope C inse	638	50	18	AAW28348	Staphylococcus aur
566	5	10.9	38	20	AAW60359	Human normal blad	639	50	21	AAG37777	Arabidopsis thalia
567	5	10.9	38	22	AAW60530	Human brain expres	640	50	22	AAU50306	Propionibacterium
568	5	10.9	38	22	AAW73186	Human bone marrow	641	50	22	AAU59594	Propionibacterium
569	5	10.9	38	22	AAW33396	Peptide #7433 enco	642	50	22	ABB23488	Protein #5487 enco
570	5	10.9	39	18	AAW27950	Staphylococcus aur	643	50	22	AAW58929	Human brain expres
571	5	10.9	39	21	AAW54265	Amino acid fragmen	644	51	17	AAW49777	Calcium channel is
572	5	10.9	39	21	AAW55696	B. subtilis ycfB p	645	51	18	AAW32466	Mycobacterium tube
573	5	10.9	39	22	AAU01783	Human secreted pro	646	51	18	AAW33339	Mycobacterium tube
574	5	10.9	40	9	AAW80076	Cytomegalovirus (C	647	51	19	AAW81701	M. tuberculosis im
575	5	10.9	40	18	AAW18235	Fibrobacter succin	648	51	19	AAW64334	Mycobacterium tube
576	5	10.9	40	21	AAW55145	Arabidopsis thalia	649	51	20	AAW39131	M. tuberculosis ES
577	5	10.9	40	22	AAW49971	Fusarium poae hyar	650	51	20	AAW38988	M. tuberculosis re
578	5	10.9	40	22	ABW43813	Peptide #11319 enc	651	51	20	AAW38451	Human secreted pro
579	5	10.9	40	22	ABW42221	Protein #4220 enco	652	51	21	AAG34437	Arabidopsis thalia
580	5	10.9	40	22	ABW26750	Human brain expres	653	51	22	ABW10501	Human ovarian and/
581	5	10.9	40	22	AAW64779	Human brain expres	654	51	22	AAW96224	Human reproductive
582	5	10.9	40	22	AAW77538	Human bone marrow	655	51	22	AAU01897	Mycobacterium tube
583	5	10.9	40	22	AAW21471	Peptide #7905 enco	656	52	21	AAU00521	Human secreted pro
584	5	10.9	40	22	AAW37722	Peptide #11759 enc	657	52	22	AAU43770	Propionibacterium
585	5	10.9	40	22	AAU01625	Human secreted pro	658	52	22	ABW16203	Human nervous syst
586	5	10.9	41	15	AAW57186	GET epitope C inse	659	52	22	AAW06584	Human foetal prote
587	5	10.9	41	21	AAW50421	Human uncoupling p	660	52	22	AAW06729	Human foetal prote
588	5	10.9	41	21	AAW27164	hPIV2 partial prot	661	53	20	AAW42792	Streptococcus pyog
589	5	10.9	41	22	ABW31202	Peptide #3853 enco	662	54	21	AAW38489	Human secreted pro
590	5	10.9	41	22	ABW36400	Peptide #3906 enco	663	54	22	AAU43821	Propionibacterium
591	5	10.9	41	22	AAW57165	Human brain expres	664	54	22	ABG14477	Novel human diago
592	5	10.9	41	22	AAW69564	Human bone marrow	665	54	22	ABW40347	Peptide #7853 enco
593	5	10.9	41	22	AAW17385	Peptide #3819 enco	666	54	22	ABW17231	Human nervous syst
594	5	10.9	41	22	AAW29804	Peptide #3941 enco	667	54	22	ABW24735	Protein #6734 enco
595	5	10.9	41	22	AAW05074	Peptide #3756 enco	668	54	22	AAW81171	Human haematologic

669	5	10.9	54	22	AAM81182	Human haematologic	742	65	22	ABB39031	Peptide #6537 enco
670	5	10.9	54	22	AAM81581	Human brain expres	743	65	22	ABB23975	Protein #5974 enco
671	5	10.9	54	22	AAM61153	Human bone marrow	744	65	22	AAM59610	Human brain expres
672	5	10.9	54	22	AAM73867	Peptide #6565 enco	745	65	22	AAM59685	Human bone marrow
673	5	10.9	54	22	AAM20131	Propionibacterium	746	65	22	AAM72183	Human bone marrow
674	5	10.9	54	22	AAM34049	Peptide #8086 enco	747	65	22	AAM72261	Peptide #5954 enco
675	5	10.9	55	22	AAU40523	Peptide #2128 enco	748	65	22	AAM19520	Peptide #6480 enco
676	5	10.9	55	22	ABB29477	Peptide #3651 enco	749	65	22	AAM32443	Peptide #6557 enco
677	5	10.9	55	22	ABB31000	Peptide #2161 enco	750	65	22	AAM32520	Human foetal prote
678	5	10.9	55	22	ABB34655	Peptide #3699 enco	751	65	22	AAM41112	Peptide #11640 enco
679	5	10.9	55	22	ABB36193	Human nervous syst	752	66	22	ABB44134	Protein #5021 enco
680	5	10.9	55	22	ABB20062	Human brain expres	753	66	22	ABB23022	Human brain expres
681	5	10.9	55	22	AAM55445	Human brain expres	754	66	22	ABB27016	Human bone marrow
682	5	10.9	55	22	AAM56974	Human bone marrow	755	66	22	AAM65157	Peptide #8196 enco
683	5	10.9	55	22	AAM60182	Human bone marrow	756	66	22	AAM77862	Peptide #12121 enco
684	5	10.9	55	22	AAM67836	Human bone marrow	757	66	22	AAM38084	Human protein kina
685	5	10.9	55	22	AAM69161	Human polypeptide	758	66	22	ABB85512	Peptide #487 enco
686	5	10.9	55	22	AAM72800	Peptide #2085 enco	759	66	22	ABB27810	Protein #454 enco
687	5	10.9	55	22	AAM10860	Peptide #2189 enco	760	67	22	ABB32981	Human cardiovascular
688	5	10.9	55	22	AAM15651	Peptide #7067 enco	761	67	22	ABB22314	Human reproductive
689	5	10.9	55	22	AAM28152	Peptide #3570 enco	762	67	22	ABB18455	Human brain expres
690	5	10.9	55	22	AAM33030	Peptide #3570 enco	763	67	22	AAU22314	Human bone marrow
691	5	10.9	55	22	AAM03390	Serine protease pf	764	67	22	AAM53779	Human polypeptide
692	5	10.9	55	22	AAM04888	Arabidopsis thalia	765	67	22	AAO11655	Human polypeptide
693	5	10.9	56	21	AAG01174	Arabidopsis thalia	766	67	22	AAO12955	Peptide #470 enco
694	5	10.9	56	21	AAG16991	Propionibacterium	767	67	22	AAO14036	Peptide #479 enco
695	5	10.9	56	21	AAG48841	Human musculoskele	768	67	22	AAO17777	Peptide #459 enco
696	5	10.9	56	21	AAG60540	Flea serine protea	769	67	22	AAO07905	Human secreted pro
697	5	10.9	56	22	ABB03535	Human immune/haema	770	67	22	ABG00305	Novel human diagno
698	5	10.9	56	22	ABB05070	Human EST encoded	771	68	22	ABG08231	Human immune/haema
699	5	10.9	57	21	AAB38625	Human EST encoded	772	68	22	AAO13010	Human polypeptide
700	5	10.9	57	22	AAM84903	Human colon cancer	773	68	22	AAO30649	Peptide #3300 enco
701	5	10.9	57	22	AAM23726	Human immune/haema	774	68	22	ABB30649	Peptide #3398 enco
702	5	10.9	57	22	AAM24492	Human polypeptide	775	68	22	ABB35820	Peptide #3430 enco
703	5	10.9	58	21	AAO13395	Linker sequence us	776	69	22	ABB21236	Protein #3235 enco
704	5	10.9	58	21	AAO13672	Staphylococcus aur	777	69	22	ABB21333	Protein #3332 enco
705	5	10.9	59	22	AAO13672	Amino acid sequenc	778	69	22	AAO56727	Human brain expres
706	5	10.9	59	22	AAO13672	Human secreted pro	779	69	22	AAO59004	Human bone marrow
707	5	10.9	59	22	AAO13672	Chemically modifie	780	69	22	AAO59105	Human immune/haema
708	5	10.9	60	16	AAO13672	Propionibacterium	781	69	22	AAO59004	Peptide #3375 enco
709	5	10.9	60	18	AAO13672	Human immune/haema	782	69	22	AAO16941	Peptide #3358 enco
710	5	10.9	60	20	AAO13672	Novel human diagno	783	69	22	AAO29321	Peptide #3463 enco
711	5	10.9	60	21	AAO13672	Human immune/haema	784	69	22	AAO29426	Peptide #3316 enco
712	5	10.9	60	22	AAO13672	Human polypeptide	785	69	22	AAO4634	Rat brain calcium
713	5	10.9	60	22	AAO13672	Human secreted pro	786	70	22	AAO97850	Human prostate can
714	5	10.9	60	22	AAO13672	Zea mays protein f	787	70	22	AAO50030	Propionibacterium
715	5	10.9	61	22	AAO13672	C glutamicum prote	788	70	22	AAO52829	Propionibacterium
716	5	10.9	61	22	AAO13672	Human presenilin I	789	70	22	AAO59695	Human 5' EST secre
717	5	10.9	61	22	AAO13672	Zea mays protein f	790	71	22	AAO11890	Pinus radiata tran
718	5	10.9	61	22	AAO13672	Zea mays protein f	791	71	22	AAO33189	Zea mays protein f
719	5	10.9	61	22	AAO13672	Arabidopsis thalia	792	71	22	AAO18763	Arabidopsis thalia
720	5	10.9	62	21	AAO13672	Human haematologic	793	71	22	AAO32668	Arabidopsis thalia
721	5	10.9	62	21	AAO13672	Human haematologic	794	71	22	AAO37983	Human musculoskele
722	5	10.9	62	21	AAO13672	Human reproductive	795	71	22	AAO48400	Human immune/haema
723	5	10.9	62	21	AAO13672	Human immune/haema	796	71	22	AAO40401	Human immune/haema
724	5	10.9	63	19	AAO13672	Human polypeptide	797	71	22	AAO83013	Human nervous syst
725	5	10.9	63	21	AAO13672	Propionibacterium	798	71	22	AAO88707	Human human secret
726	5	10.9	63	21	AAO13672	Human novel foetal	799	72	22	AAO15198	Human reproductive
727	5	10.9	63	21	AAO13672	Human EST encoded	800	72	22	AAO29770	Human breast or ov
728	5	10.9	63	22	AAO13672	Human secreted pro	801	72	22	AAO42317	Human prostate tum
729	5	10.9	63	22	AAO13672	Lactobacillus amyl	802	72	22	AAO73882	Novel human diagno
730	5	10.9	63	22	AAO13672	Human prostate can	803	73	22	AAO27287	Propionibacterium
731	5	10.9	63	22	AAO13672	Human 5' EST secre	804	73	22	AAO56270	Human reproductive
732	5	10.9	63	22	AAO13672	Peptide #6465 enco	805	74	22	AAO39326	
733	5	10.9	64	22	AAO13672		806	74	22		
734	5	10.9	64	22	AAO13672		807	74	22		
735	5	10.9	64	22	AAO13672		808	74	22		
736	5	10.9	64	22	AAO13672		809	74	22		
737	5	10.9	65	18	AAO13672		810	74	22		
738	5	10.9	65	20	AAO13672		811	74	22		
739	5	10.9	65	20	AAO13672		812	74	22		
740	5	10.9	65	22	AAO13672		813	74	22		
741	5	10.9	65	22	AAO13672		814	74	22		

815	5	10.9	74	22	AAB64402	Amino acid sequenc	888	5	10.9	84	21	AAG35114	Arabidopsis thalia
816	5	10.9	75	21	AAB08492	Amino acid sequenc	889	5	10.9	84	21	AAG36650	Arabidopsis thalia
817	5	10.9	75	21	AAG20590	Arabidopsis thalia	890	5	10.9	84	21	AAG43509	Arabidopsis thalia
818	5	10.9	75	22	AAG67855	Murine leukemia vi	891	5	10.9	84	21	AAG44608	Arabidopsis thalia
819	5	10.9	75	22	AAB87445	Human gene 42 enco	892	5	10.9	84	21	AAG46379	Arabidopsis thalia
820	5	10.9	76	18	AAB20311	H. pylori secreted	893	5	10.9	84	21	AAG46391	Arabidopsis thalia
821	5	10.9	76	21	AAB45107	Human secreted pro	894	5	10.9	84	21	AAG46415	Arabidopsis thalia
822	5	10.9	76	22	ABB14839	Human nervous syst	895	5	10.9	84	21	AAG50740	Arabidopsis thalia
823	5	10.9	76	22	AAM61365	Human brain expres	896	5	10.9	84	21	AAG50749	Arabidopsis thalia
824	5	10.9	76	22	AAW74110	Human bone marrow	897	5	10.9	84	22	AAG02176	Human secreted pro
825	5	10.9	76	22	AAW84253	Human immune/haema	898	5	10.9	84	22	AAW89076	Human immune/haema
826	5	10.9	76	22	AAW86311	Human immune/haema	899	5	10.9	84	22	AAW06597	Human foetal prote
827	5	10.9	76	22	AAW93165	Human digestive sy	900	5	10.9	84	22	AAW06793	Human foetal prote
828	5	10.9	76	22	AAW86311	Human polypeptide	901	5	10.9	84	22	AAW06793	Human foetal prote
829	5	10.9	76	22	AAW07910	Human polypeptide	902	5	10.9	84	22	AAW06793	Human foetal prote
830	5	10.9	77	18	AAW20663	Human polypeptide	903	5	10.9	85	21	AAW37314	Arabidopsis thalia
831	5	10.9	77	19	AAW75001	H. Pylori secreted	904	5	10.9	85	21	AAW48399	Arabidopsis thalia
832	5	10.9	77	19	AAW60998	Human secreted pro	905	5	10.9	85	22	AAW07257	Human polypeptide
833	5	10.9	77	21	AAW34363	Streptococcus pneu	906	5	10.9	85	22	AAW63482	Human gastric can
834	5	10.9	77	22	AAW28103	Human secreted pro	907	5	10.9	86	21	AAW50420	Human uncoupling p
835	5	10.9	77	22	AAW33278	Human peptide #754	908	5	10.9	86	21	AAW51114	Arabidopsis thalia
836	5	10.9	77	22	AAW18740	Peptide #784 enco	909	5	10.9	86	22	AAU42401	Propionibacterium
837	5	10.9	77	22	AAW54070	Protein #739 enco	910	5	10.9	86	22	AAU62283	Propionibacterium
838	5	10.9	77	22	AAW66458	Human brain expres	911	5	10.9	86	22	ABW08617	Novel human diagno
839	5	10.9	77	22	AAW70377	Human bone marrow	912	5	10.9	86	22	ABW50756	Human secreted pro
840	5	10.9	77	22	AAW14330	Human bone marrow	913	5	10.9	86	22	ABW11368	Human secreted pro
841	5	10.9	77	22	AAW26740	Peptide #764 enco	914	5	10.9	87	18	AAW18330	Bovine p32 ll-cis-
842	5	10.9	77	22	AAW26740	Peptide #777 enco	915	5	10.9	87	20	AAW35743	Chlamydia pneumoni
843	5	10.9	77	22	AAW20059	Peptide #741 enco	916	5	10.9	87	20	AAW01996	PCR clone 194, a f
844	5	10.9	78	18	AAW23001	Canine herpesvirus	917	5	10.9	87	21	AAW09639	Arabidopsis thalia
845	5	10.9	78	19	AAW72661	Canine herpes viru	918	5	10.9	87	21	AAW20589	Arabidopsis thalia
846	5	10.9	78	20	AAW06239	Human cadherin-3 p	919	5	10.9	87	21	AAW42420	Arabidopsis thalia
847	5	10.9	78	21	AAW54871	Arabidopsis thalia	920	5	10.9	87	21	AAW49867	Arabidopsis thalia
848	5	10.9	78	22	AAW82026	Human haematologic	921	5	10.9	87	21	AAW54324	Arabidopsis thalia
849	5	10.9	78	22	AAW64235	Human brain expres	922	5	10.9	87	21	AAW56339	Arabidopsis thalia
850	5	10.9	78	22	AAW53138	Canine herpes viru	923	5	10.9	87	22	AAW82901	Human immune/haema
851	5	10.9	78	22	AAW63473	Human gastric can	924	5	10.9	87	22	AAW92990	C glutamicum prote
852	5	10.9	79	21	AAW43069	Human OREF ORF2833	925	5	10.9	88	17	AAW01206	Staphylococcus aur
853	5	10.9	79	21	AAW04436	Arabidopsis thalia	926	5	10.9	88	19	AAW77527	Serine protease Pf
854	5	10.9	79	21	AAW09640	Arabidopsis thalia	927	5	10.9	88	20	AAW73898	Human prostate tum
855	5	10.9	79	21	AAW42421	Arabidopsis thalia	928	5	10.9	88	21	AAW37313	Arabidopsis thalia
856	5	10.9	79	21	AAW49868	Arabidopsis thalia	929	5	10.9	88	21	AAW41556	Arabidopsis thalia
857	5	10.9	79	22	AAW03620	Human secreted pro	930	5	10.9	88	22	AAU21374	Human novel foetal
858	5	10.9	79	22	AAW32616	Novel human secret	931	5	10.9	88	22	AAW82687	Human immune/haema
859	5	10.9	79	22	AAW03643	Human polypeptide	932	5	10.9	88	22	AAW06338	Human foetal prote
860	5	10.9	80	20	AAW89911	Antigen 2 from clu	933	5	10.9	88	22	AAW632383	Shiga toxin beta s
861	5	10.9	80	21	AAW02452	Human secreted pro	934	5	10.9	88	22	AAW61901	Mouse apoptosis me
862	5	10.9	80	22	AAW69642	Drosophila melanog	935	5	10.9	88	22	AAW50603	Flea serine protea
863	5	10.9	80	22	AAW23267	Novel human enzyme	936	5	10.9	89	13	AAW22429	Murine CP-10. Mus
864	5	10.9	80	22	AAW010298	Human polypeptide	937	5	10.9	89	17	AAW06404	Verotoxinigenic E. c
865	5	10.9	80	22	AAW63490	Human gastric can	938	5	10.9	89	19	AAW58828	Shiga toxin type 1
866	5	10.9	81	22	AAW42097	Human polypeptide	939	5	10.9	89	20	AAW59968	Human endometrium
867	5	10.9	81	22	AAW67375	Propionibacterium	940	5	10.9	89	20	AAW29338	Human secreted pro
868	5	10.9	82	13	AAW27371	TRFP Chain #2 C2ST	941	5	10.9	89	21	AAW41006	Human OREF ORF770
869	5	10.9	82	13	AAW27371	Arabidopsis thalia	942	5	10.9	89	21	AAW34223	Arabidopsis thalia
870	5	10.9	82	22	AAW23267	Novel human secret	943	5	10.9	89	21	AAW36152	Arabidopsis thalia
871	5	10.9	82	22	AAW05273	Human polypeptide	944	5	10.9	89	21	AAW96682	E. coli verotoxin
872	5	10.9	82	22	AAW3486	Human EST encoded	945	5	10.9	89	22	AAW44594	Mouse wound healin
873	5	10.9	82	18	AAW27792	Staphylococcus aur	946	5	10.9	89	22	AAU39089	Human secreted pro
874	5	10.9	83	21	AAW40416	Human OREF ORF180	947	5	10.9	89	22	AAW05291	Human polypeptide
875	5	10.9	83	21	AAW24737	Plant SDF encoded	948	5	10.9	89	22	AAW11912	Human polypeptide
876	5	10.9	83	21	AAW51439	Arabidopsis thalia	949	5	10.9	89	22	AAW60523	Bovine ghrelin pre
877	5	10.9	83	21	AAW51444	Arabidopsis thalia	950	5	10.9	89	22	AAW42419	Arabidopsis thalia
878	5	10.9	83	21	AAW51450	Arabidopsis thalia	951	5	10.9	90	21	AAW02990	Human secreted pro
879	5	10.9	83	21	AAW51457	Arabidopsis thalia	952	5	10.9	90	22	AAW96654	Human reproductive
880	5	10.9	83	21	AAW51457	Arabidopsis thalia	953	5	10.9	90	22	AAW84788	Human immune/haema
881	5	10.9	83	22	AAW89159	Human immune/haema	954	5	10.9	90	22	AAU14264	Human novel protei
882	5	10.9	83	22	AAW89175	Human immune/haema	955	5	10.9	90	22	AAU14500	Human novel protei
883	5	10.9	84	21	AAW06190	Arabidopsis thalia	956	5	10.9	91	20	AAU14480	Fragment of human
884	5	10.9	84	21	AAW11537	Arabidopsis thalia	957	5	10.9	91	21	AAW18762	Zea mays protein f
885	5	10.9	84	21	AAW26576	Arabidopsis thalia	958	5	10.9	91	21	AAW37981	Arabidopsis thalia
886	5	10.9	84	21	AAW26603	Arabidopsis thalia	959	5	10.9	91	21	AAW00799	Human secreted pro
887	5	10.9	84	21	AAW27871	Arabidopsis thalia	960	5	10.9	91	21	AAW00800	Human secreted pro

```

961 5 10.9 91 22 ABG21730 Novel human diagno
962 5 10.9 91 22 AB503894 Human musculoskele
963 5 10.9 91 22 AAM85962 Human immune/haema
964 5 10.9 91 22 AAB63492 Human gastric canc
965 5 10.9 92 15 AAR65454 T-cell receptor v-
966 5 10.9 92 20 AAY37880 Chlamydia trachoma
967 5 10.9 92 21 AAB09885 Hsp70 C-terminal 9
968 5 10.9 92 21 AAG25229 Arabidopsis thalia
969 5 10.9 92 21 AAG34610 Arabidopsis thalia
970 5 10.9 92 21 AAG01138 Human secreted pro
971 5 10.9 92 22 AAU48287 Propionibacterium
972 5 10.9 92 22 AAG08059 Novel human diagno
973 5 10.9 92 22 AAO00568 Human polypeptide
974 5 10.9 93 21 AAG24552 Arabidopsis thalia
975 5 10.9 93 21 AAG56414 Arabidopsis thalia
976 5 10.9 93 22 AAU36001 Helicobacter pylor
977 5 10.9 93 22 AAU56795 Propionibacterium
978 5 10.9 93 22 ABG19225 Novel human diagno
979 5 10.9 94 19 AAW79096 Human secreted pro
980 5 10.9 94 21 AAB57028 Human prostate can
981 5 10.9 94 21 AAG36543 Arabidopsis thalia
982 5 10.9 94 21 AAG55575 Arabidopsis thalia
983 5 10.9 94 22 ABG28213 Novel human diagno
984 5 10.9 94 22 AAU31580 Novel human secret
985 5 10.9 95 16 AAW11494 ESAT6. Mycobacter
986 5 10.9 95 18 AAW28022 Amino acid sequenc
987 5 10.9 95 19 AAW67637 MSRV-1 virus clone
988 5 10.9 95 19 AAW85049 Amino acid sequenc
989 5 10.9 95 20 AAY66143 Human bladder tumo
990 5 10.9 95 20 AAY59742 Human normal ovar
991 5 10.9 95 20 AAY29890 Mycobacterium tube
992 5 10.9 95 20 AAY29788 Mycobacterial tube
993 5 10.9 95 21 AAB44119 Human cancer assoc
994 5 10.9 95 21 AAG25735 Arabidopsis thalia
995 5 10.9 95 21 AAG37927 Arabidopsis thalia
996 5 10.9 95 21 AAG60424 Arabidopsis thalia
997 5 10.9 95 22 AAB35219 M tuberculosis RV3
998 5 10.9 96 21 AAB40905 Human ORFX ORF669
999 5 10.9 96 22 ABG10913 Novel human diagno
1000 5 10.9 97 12 AAR12123 TRFP chain 2 - tru

ALIGNMENTS

RESULT 1
ID AAO00376 standard; Protein; 115 AA.
XX
AC AAO00376;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 14268.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX

Tang YT, Liu C, Drmanac RT;
WPI: 2001-514838/56.
N-PSDB; AAI80307.
Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukaemia, inflammation and immune
disorders -
Claim 20; SEQ ID NO 14268; 1399pp + Sequence Listing; English.
The invention relates to human polynucleotides (AAI79941-AAI93841) and
the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 115 AA;
Query Match 15.2%; Score 7; DB 22; Length 115;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 IIIGGSG 27
| | | | |
Db 28 IIIGGSG 34
RESULT 2
AAO01734
ID AAO01734 standard; Protein; 118 AA.
XX
AC AAO01734;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 15626.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
Tang YT, Liu C, Drmanac RT;
WPI: 2001-514838/56.
N-PSDB; AAI81665.
Isolated nucleic acids and polypeptides, useful for preventing
diagnosing and treating e.g. leukaemia, inflammation and immune
disorders -
```

XX PS Claim 20; SEQ ID NO 15626; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 118 AA;

Query Match 15.2%; Score 7; DB 22; Length 118;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SLKNDK 12

DB 108 SLKNDK 114

|||||

RESULT 3

AAO00422

ID AAO00422 standard; Protein; 126 AA.

XX AC AAO00422;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 14314.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX N-PSDB; AAI80353.

XX PT Isolated nucleic acids and polypeptides, useful for preventing

XX PT diagnosing and treating e.g. leukaemia, inflammation and immune

XX PT disorders -

XX PS Claim 20; SEQ ID NO 14314; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published_pct_sequences.

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 126 AA;

Query Match 15.2%; Score 7; DB 22; Length 126;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 IIIGSG 27

DB 67 IIIGSG 73

|||||

RESULT 4

AAO11182

ID AAO11182 standard; Protein; 150 AA.

XX AC AAO11182;

XX DT 06-NOV-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 25074.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;

XX KW nervous system disorders; arthritis; inflammation.

XX OS Homo sapiens.

XX PN WO200164835-A2.

XX PD 07-SEP-2001.

XX PF 26-FEB-2001; 2001WO-US04927.

XX PR 28-FEB-2000; 2000US-0515126.

XX PR 18-MAY-2000; 2000US-0577409.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT;

XX WPI: 2001-514838/56.

XX N-PSDB; AAI91113.

XX PT Isolated nucleic acids and polypeptides, useful for preventing

XX PT diagnosing and treating e.g. leukaemia, inflammation and immune

XX PT disorders -

XX PS Claim 20; SEQ ID NO 25074; 1399pp + Sequence Listing; English.

XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and

XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to

XX CC cytokine, cell proliferation or cell differentiation or which may induce

XX CC production of other cytokines in other cell populations. The

XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX CC peptide therapy. The polypeptides have various cytokine-like activities,

XX CC e.g. stem cell growth factor activity, haematopoiesis regulating

XX CC activity, tissue growth factor activity, immunomodulatory activity and

XX CC activin/inhibin activity and may be useful in the diagnosis and/or

XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX CC inflammation.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

```
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 150 AA;
    Query Match 15.2%; Score 7; DB 22; Length 150;
    Best Local Similarity 100.0%; Pred. No. 25;
    Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 IIIGSG 27
Db 68 IIIGSG 74

RESULT 5
AAG19904
ID AAG19904 standard; Protein; 186 AA.
XX
AC AAG19904;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 21882.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145275.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
```

PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0156659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159285.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160988.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 15.2%; Score 7; DB 21; Length 186;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 25 GSGSLST 31
 Db 17 GSGSLST 23
 RESULT 6
 AAG19903
 ID AAG19903 standard; Protein; 257 AA.
 XX AAG19903;
 AC AAG19903;
 XX 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 21881.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 PN EPI033405-A2.
 PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 25-MAY-1999; 99US-0135629.
 PR 27-MAY-1999; 99US-0136021.
 PR 28-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 03-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.

us-09-833-017b-2.oligo.rag

Wed Nov 13 14:19:03 2002

PR	17-JUN-1999;	99US-0139492.	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	99US-0139454.	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	99US-0139455.	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	99US-0139456.	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	99US-0139457.	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139458.	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	99US-0139459.	PR	20-AUG-1999;	99US-0149923.
PR	18-JUN-1999;	99US-0139460.	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	99US-0139461.	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	99US-0139462.	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139463.	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139464.	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	99US-0139465.	PR	27-AUG-1999;	99US-0151066.
PR	18-JUN-1999;	99US-0139466.	PR	27-AUG-1999;	99US-0151080.
PR	21-JUN-1999;	99US-0139763.	PR	30-AUG-1999;	99US-0151303.
PR	22-JUN-1999;	99US-0139899.	PR	31-AUG-1999;	99US-0151438.
PR	23-JUN-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151930.
PR	23-JUN-1999;	99US-0140354.	PR	07-SEP-1999;	99US-0152363.
PR	24-JUN-1999;	99US-0140695.	PR	10-SEP-1999;	99US-0153070.
PR	28-JUN-1999;	99US-0140823.	PR	13-SEP-1999;	99US-0153758.
PR	29-JUN-1999;	99US-0140991.	PR	15-SEP-1999;	99US-0154018.
PR	30-JUN-1999;	99US-0141287.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0141842.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155139.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155486.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159637.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0144814.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145086.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145145.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160981.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160989.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0161404.
PR	27-JUL-1999;	99US-0145313.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145318.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145319.	PR	25-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161360.
PR	28-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	28-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.			
PR	05-AUG-1999;	99US-0147192.			
PR	06-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			

Query Match 15.2%; Score 7; DB 21; Length 257;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSLST 31
 Db 88 GSGSLST 94

RESULT 7

AAG19902
ID AAG19902 standard; Protein; 287 AA.
XX
AC AAG19902;
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 21880.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 18-JUN-1999; 99US-0139460.
XX 05-MAR-1999; 99US-0123180.
PR 18-JUN-1999; 99US-0139461.
XX 09-MAR-1999; 99US-0123548.
PR 18-JUN-1999; 99US-0139462.
XX 23-MAR-1999; 99US-0125788.
PR 18-JUN-1999; 99US-0139463.
XX 25-MAR-1999; 99US-0126264.
PR 18-JUN-1999; 99US-0139750.
XX 29-MAR-1999; 99US-0126785.
PR 18-JUN-1999; 99US-0139763.
XX 01-APR-1999; 99US-0127462.
PR 21-JUN-1999; 99US-0139817.
XX 06-APR-1999; 99US-0128234.
PR 22-JUN-1999; 99US-0139899.
XX 08-APR-1999; 99US-0128714.
PR 23-JUN-1999; 99US-0140353.
XX 10-APR-1999; 99US-0129845.
PR 24-JUN-1999; 99US-0140354.
XX 16-APR-1999; 99US-0130077.
PR 24-JUN-1999; 99US-0140695.
XX 21-APR-1999; 99US-0130449.
PR 28-JUN-1999; 99US-0140823.
XX 23-APR-1999; 99US-0130891.
PR 29-JUN-1999; 99US-0140991.
XX 30-APR-1999; 99US-0131449.
PR 30-JUN-1999; 99US-0141287.
XX 01-MAY-1999; 99US-0132407.
PR 01-JUL-1999; 99US-0141842.
XX 04-MAY-1999; 99US-0132484.
PR 02-JUL-1999; 99US-0142055.
XX 08-MAY-1999; 99US-0132488.
PR 06-JUL-1999; 99US-0142390.
XX 10-MAY-1999; 99US-0132488.
PR 08-JUL-1999; 99US-0142803.
XX 16-MAY-1999; 99US-0132488.
PR 12-JUL-1999; 99US-0142920.
XX 22-MAY-1999; 99US-0132488.
PR 14-JUL-1999; 99US-0143542.
XX 28-MAY-1999; 99US-0132488.
PR 15-JUL-1999; 99US-0143624.
XX 06-JUN-1999; 99US-0132488.
PR 16-JUL-1999; 99US-0144005.
XX 13-JUN-1999; 99US-0143624.
PR 19-JUL-1999; 99US-0144086.
XX 20-JUN-1999; 99US-0144325.
PR 15-JUL-1999; 99US-0144332.
XX 27-JUN-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
XX 29-JUN-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
XX 30-JUN-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
XX 01-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
XX 02-JUL-1999; 99US-0144884.
PR 20-JUL-1999; 99US-0144884.
XX 08-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
XX 19-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
XX 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
XX 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145192.
XX 24-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
XX 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
XX 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
XX 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
XX 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146386.
XX 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
XX 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
XX 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
XX 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
XX 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
XX 10-AUG-1999; 99US-0148171.
PR 10-AUG-1999; 99US-0148319.
XX 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
XX 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0148684.
XX 16-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149368.
XX 17-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
XX 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149426.
XX 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149722.
XX 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150556.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151086.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 15.2%; Score 7; DB 21; Length 287;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSLST 31
 Db 118 GSGSLST 124

RESULT 8
 AAR42086
 ID AAR42086 standard; protein: 558 AA.
 AC AAR42086;
 XX AC
 DT 05-MAY-1994 (first entry)
 XX

DE Mouse RelB protein.
 XX NF-KappaB/Rel/Dorsal; NRD; DNA binding region.
 KW Mus musculus.
 OS W09320219-A.
 PN 14-OCT-1993.
 PD 01-APR-1993; 93WO-US03027.
 PF 06-APR-1992; 92US-0862987.
 PR (USSH) US SEC DEPT HEALTH.
 PA Leonard WJ, Toledano MB;
 PI WPI; 1993-336925/42.
 DR Mutation of DNA binding region of NF-KB/rel/dorsal protein - to
 PT prevent or control binding to DNA for control of gene activation
 PS Disclosure; Page 25-26; 52pp; English.
 XX The sequence is that of the mouse RelB protein, the DNA binding
 CC region of which may be mutated by replacing one or more amino acids.
 CC This allows inhibition, prevention or control of DNA binding and
 CC thus activation of genes controlled by the protein can be controlled.
 XX SQ Sequence 558 AA;
 Query Match 15.2%; Score 7; DB 14; Length 558;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 16 TDELEII 22
 Db 33 TDELEII 39
 RESULT 9
 AAG30638
 ID AAG30638 standard; Protein: 586 AA.
 XX AC AAG30638;
 XX DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 36664.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 XX EP1033405-A2.
 XX 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-APR-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151086.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	14-OCT-1999;	99US-0159330.	PR	04-MAY-1999;	99US-0132484.
PR	14-OCT-1999;	99US-0159331.	PR	05-MAY-1999;	99US-0132485.
PR	14-OCT-1999;	99US-0159637.	PR	06-MAY-1999;	99US-0132486.
PR	14-OCT-1999;	99US-0159638.	PR	06-MAY-1999;	99US-0132487.
PR	18-OCT-1999;	99US-0159584.	PR	07-MAY-1999;	99US-0132863.
PR	21-OCT-1999;	99US-0160741.	PR	11-MAY-1999;	99US-0134256.
PR	21-OCT-1999;	99US-0160767.	PR	14-MAY-1999;	99US-0134218.
PR	21-OCT-1999;	99US-0160768.	PR	14-MAY-1999;	99US-0134219.
PR	21-OCT-1999;	99US-0160770.	PR	14-MAY-1999;	99US-0134221.
PR	21-OCT-1999;	99US-0160814.	PR	14-MAY-1999;	99US-0134370.
PR	21-OCT-1999;	99US-0160815.	PR	18-MAY-1999;	99US-0134768.
PR	22-OCT-1999;	99US-0160980.	PR	19-MAY-1999;	99US-0134941.
PR	22-OCT-1999;	99US-0160981.	PR	20-MAY-1999;	99US-0135124.
PR	22-OCT-1999;	99US-0160989.	PR	21-MAY-1999;	99US-0135353.
PR	25-OCT-1999;	99US-0161404.	PR	21-MAY-1999;	99US-0135629.
PR	25-OCT-1999;	99US-0161405.	PR	24-MAY-1999;	99US-0136021.
PR	25-OCT-1999;	99US-0161406.	PR	25-MAY-1999;	99US-0136392.
PR	26-OCT-1999;	99US-0161359.	PR	27-MAY-1999;	99US-0136782.
PR	26-OCT-1999;	99US-0161360.	PR	28-MAY-1999;	99US-0137222.
PR	26-OCT-1999;	99US-0161361.	PR	01-JUN-1999;	99US-0137528.
PR	28-OCT-1999;	99US-0161320.	PR	03-JUN-1999;	99US-0137502.
PR	28-OCT-1999;	99US-0161992.	PR	04-JUN-1999;	99US-0137724.
PR	28-OCT-1999;	99US-0161993.	PR	07-JUN-1999;	99US-0138094.
PR	29-OCT-1999;	99US-0162142.	PR	08-JUN-1999;	99US-0138540.
Query Match 15.2%; Score 7; DB 21; Length 586;					99US-0138847.
Best Local Similarity 100.0%; Pred. No. 89;					99US-0139119.
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					99US-0139452.
Oy	21 IIIIGSG 27				99US-0139453.
					99US-0139492.
Db	423 IIIIGSG 429				99US-0139454.
RESULT 10					
AAG30637					
ID	AAG30637 standard; Protein; 612 AA.				99US-0139455.
XX	AC AAG30637;				99US-0139456.
XX	DF 17-OCT-2000 (first entry)				99US-0139457.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 36663.				99US-0139458.
XX	protein identification; signal transduction pathway; metabolic pathway;				99US-0139459.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;				99US-0139460.
KW	termination sequence.				99US-0139461.
XX	OS Arabidopsis thaliana.				99US-0139462.
XX	PN EP1033405-A2.				99US-0139463.
XX	PD 06-SEP-2000.				99US-0139750.
XX	PF 25-FEB-2000; 2000EP-0301439.				99US-0139763.
XX	25-FEB-1999;	99US-0121825.			99US-0139817.
PR	05-MAR-1999;	99US-0123180.			99US-0139899.
PR	09-MAR-1999;	99US-0123548.			99US-0140353.
PR	23-MAR-1999;	99US-0125788.			99US-0140354.
PR	23-MAR-1999;	99US-0126264.			99US-0140695.
PR	29-MAR-1999;	99US-0126785.			99US-0140823.
PR	01-APR-1999;	99US-0127462.			99US-0140991.
PR	06-APR-1999;	99US-0128234.			99US-0141287.
PR	08-APR-1999;	99US-0128714.			99US-0141842.
PR	16-APR-1999;	99US-0129845.			99US-0142154.
PR	19-APR-1999;	99US-0130077.			99US-0142055.
PR	21-APR-1999;	99US-0130449.			99US-0142390.
PR	23-APR-1999;	99US-0130510.			99US-0142803.
PR	23-APR-1999;	99US-0130891.			99US-0142920.
PR	28-APR-1999;	99US-0131449.			99US-0142977.
PR	30-APR-1999;	99US-0132048.			99US-0143542.
PR	30-APR-1999;	99US-0132407.			99US-0143624.

```
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145911.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147410.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 15.2%; Score 7; DB 21; Length 612;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 I I I G S G 27
Db 449 I I I G S G 455

RESULT 11
AAG30636
ID AAG30636 standard; Protein; 618 AA.
XX
AC AAG30636;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 36662.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
```

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140354.
PR 28-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143342.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.

PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 15.2%; Score 7; DB 21; Length 618;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 IIGSG 27
 |||||
 Db 455 IIGSG 461

RESULT 12
 ABG17981
 ID ABG17981 standard; Protein; 658 AA.
 XX AC
 XX ABG17981;
 XX AC
 DT 18-FEB-2002 (first entry)
 XX DE
 DE Novel human diagnostic protein #17972.
 XX KW
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS
 OS Homo sapiens.
 XX PN
 PN WO200175067-A2.
 XX PD
 PD 11-OCT-2001.
 XX PF
 PF 30-MAR-2001; 2001WO-US08631.
 XX PR
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX PA
 PA (HYSE-) HYSEQ INC.
 XX PI
 PI Drmanac RT, Liu C, Tang YT;
 XX DR
 DR N-PSDB; AAS82168.
 XX PT
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX PS
 PS Claim 20; SEQ ID No 48340; 103pp; English.
 XX CC
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 658 AA;

Query Match 15.2%; Score 7; DB 22; Length 658;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 DELEIII 23
 |||||
 Db 378 DELEIII 384

RESULT 13
 ABB58240
 ID ABB58240 standard; Protein; 767 AA.
 XX AC
 XX ABB58240;
 XX AC
 DT 26-MAR-2002 (first entry)
 XX DE
 DE Drosophila melanogaster polypeptide SEQ ID NO 1512.

XX KW
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX OS
 OS Drosophila melanogaster.
 XX PN
 PN WO200171042-A2.
 XX PD
 PD 27-SEP-2001.
 XX PF
 PF 23-MAR-2001; 2001WO-US09231.
 XX PR
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX PA
 PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL02343.
 XX PT
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX PS
 PS Disclosure; SEQ ID NO 1512; 21pp + Sequence Listing; English.

XX CC
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 767 AA;

Query Match 15.2%; Score 7; DB 22; Length 767;
 Best Local Similarity 100.0%; Pred. No. 11e-02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Drosophila melanogaster polypeptide SEQ ID NO 18432.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL07983.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 18432; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2771 AA;
Query Match 15.2%; Score 7; DB 22; Length 2771;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 24 GSGSLS 30
DB 1898 GSGSLS 1904
Search completed: November 5, 2002, 11:04:58
Job time : 78.3881 secs

QY 24 GSGSLS 30
DB 304 GSGSLS 310
RESULT 14
ABB70857
ID ABB70857 standard; Protein; 929 AA.
XX
AC ABB70857;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 39363.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR N-PSDB; ABL4960.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 39363; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 929 AA;
Query Match 15.2%; Score 7; DB 22; Length 929;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 39 SFTQALG 45
DB 103 SFTQALG 109
RESULT 15
ABB63880
ID ABB63880 standard; Protein; 2771 AA.
XX
AC ABB63880;
XX
XX 26-MAR-2002 (first entry)
DT
XX

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 11:03:27 ; Search time 15.1045 Seconds
(without alignments)
74.387 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 46

Sequence: I MKKTLKNDKFKETDELE.....GSLTFFRLFNKSFQALGK 46

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/pctus_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	7	15.2	558	5	PCT-US93-03027-2
2	7	15.2	767	4	US-08-836-567-8
3	6	13.0	17	1	US-08-370-567-18
4	6	13.0	17	1	US-08-438-759-18
5	6	13.0	17	5	PCT-US94-05684-18
6	6	13.0	114	2	US-08-652-558-52
7	6	13.0	134	4	US-09-367-953B-120
8	6	13.0	134	4	US-09-367-953B-121
9	6	13.0	282	4	US-08-482-918-56
10	6	13.0	282	4	US-09-224-681-56
11	6	13.0	282	4	US-08-336-728A-56
12	6	13.0	322	4	US-09-383-586-33
13	6	13.0	397	3	US-09-079-415-2
14	6	13.0	398	1	US-08-328-314-2
15	6	13.0	398	1	US-08-731-045-2
16	6	13.0	412	1	US-08-208-007A-12
17	6	13.0	412	4	US-08-974-691-4
18	6	13.0	439	4	US-09-383-586-32
19	6	13.0	446	4	US-09-268-992-51
20	6	13.0	466	4	US-09-288-992-64
21	6	13.0	477	4	US-09-446-504-19
22	6	13.0	477	4	US-09-712-266-19
23	6	13.0	477	4	US-09-268-992-4
24	6	13.0	493	4	US-09-177-349-5
25	6	13.0	495	4	US-09-268-992-2
26	6	13.0	499	2	US-07-952-853-6
27	6	13.0	499	2	US-08-914-848-6

508	2	US-08-818-024-3	Sequence 3, Appli
508	4	US-09-334-775A-3	Sequence 3, Appli
529	4	US-09-383-586-31	Sequence 31, Appl
534	2	US-08-691-814B-8	Sequence 8, Appli
538	2	US-08-541-759B-2	Sequence 2, Appli
547	2	US-08-467-822-35	Sequence 35, Appl
547	4	US-08-432-697-35	Sequence 35, Appl
547	4	US-08-466-248-35	Sequence 35, Appl
573	2	US-08-706-209-1	Sequence 1, Appli
573	3	US-08-981-787-1	Sequence 1, Appli
573	4	US-08-461-722-1	Sequence 1, Appli
573	4	US-08-336-251-1	Sequence 1, Appli
573	5	PCT-US94-06362-1	Sequence 1, Appli
573	5	PCT-US96-11373-1	Sequence 1, Appli
573	5	PCT-US96-11375-1	Sequence 1, Appli
635	2	US-09-014-969-11	Sequence 11, Appl
719	4	US-09-386-607-2	Sequence 2, Appli
887	1	US-07-596-467-2	Sequence 2, Appli
887	1	US-07-934-374-2	Sequence 2, Appli
887	1	US-07-783-861C-4	Sequence 4, Appli
888	4	US-09-268-140-4	Sequence 4, Appli
893	1	US-07-977-434-4	Sequence 4, Appli
893	1	US-08-458-819-4	Sequence 4, Appli
893	4	US-09-105-697-10	Sequence 10, Appl
893	5	PCT-US91-07035-4	Sequence 4, Appli
913	3	US-08-911-853-17	Sequence 17, Appl
913	4	US-09-479-409-17	Sequence 17, Appl
913	4	US-09-479-453-17	Sequence 17, Appl
1968	1	US-07-745-206A-7	Sequence 7, Appli
1968	1	US-08-455-543A-45	Sequence 45, Appl
1968	2	US-08-223-305C-45	Sequence 45, Appl
1968	2	US-08-311-363-7	Sequence 7, Appli
2161	1	US-07-745-206A-2	Sequence 2, Appli
2161	1	US-08-455-543A-49	Sequence 49, Appl
2161	1	US-08-455-543A-51	Sequence 51, Appl
2161	2	US-08-223-305C-49	Sequence 49, Appl
2161	2	US-08-223-305C-51	Sequence 51, Appl
2161	2	US-08-311-363-2	Sequence 2, Appli
2161	8	US-08-148-910-11	Sequence 11, Appl
2161	8	US-08-448-937A-11	Sequence 11, Appl
10.9	8	US-09-139-802-4	Sequence 4, Appli
10.9	14	US-08-849-488-16	Sequence 16, Appl
10.9	15	US-07-664-989B-11	Sequence 11, Appl
10.9	15	US-07-664-989B-18	Sequence 18, Appl
10.9	15	US-08-726-464B-30	Sequence 30, Appl
10.9	15	US-08-933-983-74	Sequence 74, Appl
10.9	16	US-08-480-190-261	Sequence 261, App
10.9	16	US-08-474-686A-7	Sequence 7, Appli
10.9	16	US-08-488-379-261	Sequence 8, Appli
10.9	16	US-08-488-379-261	Sequence 261, App
10.9	16	PCT-US93-07545-261	Sequence 261, App
10.9	19	US-08-474-696A-1	Sequence 1, Appli
10.9	19	US-08-383-753-82	Sequence 82, Appl
10.9	23	US-08-586-772-82	Sequence 82, Appl
10.9	23	US-08-959-512-82	Sequence 82, Appl
10.9	23	US-09-512-983-82	Sequence 82, Appl
10.9	26	US-08-524-677-13	Sequence 13, Appl
10.9	29	US-09-227-357-131	Sequence 371, App
10.9	31	US-08-190-130-51	Sequence 5, Appli
10.9	31	US-09-177-249-21	Sequence 21, Appl
10.9	34	US-08-491-954-36	Sequence 36, Appl
10.9	35	US-08-491-954-39	Sequence 39, Appl
10.9	38	US-07-987-272A-18	Sequence 18, Appl
10.9	38	US-08-491-954-37	Sequence 37, Appl
10.9	41	US-08-491-954-36	Sequence 38, Appl
10.9	45	US-08-462-128-39	Sequence 39, Appl
10.9	45	US-08-463-180-39	Sequence 39, Appl
10.9	49	US-08-469-318-194	Sequence 194, App
10.9	49	US-08-468-609A-194	Sequence 194, App
10.9	49	PCT-US95-01185-194	Sequence 194, App
10.9	51	US-08-818-112-104	Sequence 104, App
10.9	51	US-08-818-111-99	Sequence 99, Appl

101	5	10.9	51	4	US-08-056-556-104	Sequence 104, App	174	1	US-08-253-155A-35	Sequence 35, Appl
102	5	10.9	52	1	US-08-346-843-8	Sequence 8, Appl	175	127	US-08-789-333F-43	Sequence 43, Appl
103	5	10.9	52	2	US-08-493-284A-8	Sequence 8, Appl	176	130	US-09-014-574-1	Sequence 1, Appl
104	5	10.9	56	1	US-08-485-455D-35	Sequence 35, Appl	177	131	US-09-553-498-4	Sequence 4, Appl
105	5	10.9	56	2	US-08-482-130C-35	Sequence 35, Appl	178	153	US-08-387-942C-49	Sequence 49, Appl
106	5	10.9	56	2	US-08-484-211C-35	Sequence 35, Appl	179	154	US-09-247-155-110	Sequence 110, App
107	5	10.9	56	3	US-08-906-769-35	Sequence 35, Appl	180	156	US-08-074-121-5	Sequence 5, Appl
108	5	10.9	56	3	US-08-906-616-35	Sequence 35, Appl	181	156	US-09-181-183-2	Sequence 2, Appl
109	5	10.9	56	4	US-08-817-795-35	Sequence 35, Appl	182	156	US-08-858-207A-334	Sequence 334, App
110	5	10.9	56	4	US-08-485-443B-35	Sequence 35, Appl	183	156	US-09-277-700-2	Sequence 2, Appl
111	5	10.9	56	4	US-08-639-075A-35	Sequence 35, Appl	184	156	PCT-US94-08447-5	Sequence 5, Appl
112	5	10.9	56	4	US-09-012-431-35	Sequence 35, Appl	185	157	US-09-025-769B-372	Sequence 372, App
113	5	10.9	56	4	US-09-012-692-35	Sequence 35, Appl	186	157	US-09-025-769B-373	Sequence 373, App
114	5	10.9	56	4	US-08-906-613-35	Sequence 35, Appl	187	159	US-08-772-270A-7	Sequence 7, Appl
115	5	10.9	56	5	PCT-US95-1442A-35	Sequence 35, App	188	159	US-09-027-449-53	Sequence 53, Appl
116	5	10.9	60	3	US-08-469-318-195	Sequence 195, App	189	159	US-08-804-444A-53	Sequence 53, Appl
117	5	10.9	60	5	US-08-468-609A-195	Sequence 195, App	190	163	US-09-026-985-53	Sequence 53, Appl
118	5	10.9	60	5	PCT-US95-01185-195	Sequence 195, App	191	163	US-09-025-769B-282	Sequence 282, App
119	5	10.9	73	2	US-08-245-511-36	Sequence 36, App	192	165	US-09-087-465-30	Sequence 30, Appl
120	5	10.9	73	2	US-08-600-993A-36	Sequence 36, App	193	169	US-08-720-625-4	Sequence 4, Appl
121	5	10.9	73	3	US-09-042-012-8	Sequence 8, Appl	194	170	US-08-764-563-5	Sequence 5, Appl
122	5	10.9	75	4	US-09-305-086-2	Sequence 8, Appl	195	172	US-08-557-309B-39	Sequence 39, Appl
123	5	10.9	75	4	US-09-457-324-8	Sequence 2, Appl	196	172	US-08-834-306-39	Sequence 39, Appl
124	5	10.9	76	1	US-07-987-272A-17	Sequence 17, Appl	197	172	US-08-993-674A-39	Sequence 39, Appl
125	5	10.9	78	1	US-08-680-726A-60	Sequence 60, Appl	198	172	US-08-789-333F-42	Sequence 42, Appl
126	5	10.9	78	4	US-09-092-409-60	Sequence 60, Appl	199	177	US-08-789-333F-44	Sequence 44, Appl
127	5	10.9	80	2	US-08-691-814B-29	Sequence 29, Appl	200	178	US-08-044-621D-32	Sequence 32, Appl
128	5	10.9	87	2	US-08-562-114B-24	Sequence 24, Appl	201	178	US-08-709-912-15	Sequence 15, Appl
129	5	10.9	87	4	US-08-729-594A-24	Sequence 24, Appl	202	178	US-09-047-370-15	Sequence 15, Appl
130	5	10.9	88	1	US-07-987-272A-1	Sequence 1, Appl	203	179	US-08-764-563-4	Sequence 4, Appl
131	5	10.9	88	3	US-08-906-769-95	Sequence 95, Appl	204	179	US-08-862-124-2	Sequence 2, Appl
132	5	10.9	88	3	US-08-906-616-95	Sequence 95, Appl	205	180	US-08-328-254-8	Sequence 8, Appl
133	5	10.9	88	4	US-08-817-795-95	Sequence 95, Appl	206	187	US-09-230-637-35	Sequence 35, Appl
134	5	10.9	88	4	US-08-639-075A-95	Sequence 95, Appl	207	191	US-08-764-563-1	Sequence 1, Appl
135	5	10.9	88	4	US-09-012-431-95	Sequence 95, Appl	208	191	US-08-720-625-2	Sequence 2, Appl
136	5	10.9	88	4	US-09-012-692-95	Sequence 95, Appl	209	196	US-08-685-992-1	Sequence 1, Appl
137	5	10.9	88	4	US-08-906-613-95	Sequence 95, Appl	210	196	US-09-144-925-1	Sequence 1, Appl
138	5	10.9	88	4	US-08-936-165A-287	Sequence 287, App	211	196	US-08-484-126-7	Sequence 7, Appl
139	5	10.9	88	5	PCT-US95-1442A-95	Sequence 95, App	212	199	US-08-602-262-2	Sequence 2, Appl
140	5	10.9	89	1	US-08-816-977A-14	Sequence 14, Appl	213	199	US-09-094-716-2	Sequence 2, Appl
141	5	10.9	89	3	US-08-816-977-4	Sequence 4, Appl	214	200	5189147-8	Patent No. 5189147
142	5	10.9	95	2	US-08-465-640-2	Sequence 2, Appl	215	202	US-09-342-084-10	Sequence 10, Appl
143	5	10.9	97	3	US-08-300-928C-10	Sequence 10, Appl	216	204	US-08-808-550-32	Sequence 32, Appl
144	5	10.9	97	3	US-08-430-944D-10	Sequence 10, Appl	217	211	US-08-276-852-34	Sequence 34, Appl
145	5	10.9	97	3	US-08-430-014-10	Sequence 10, Appl	218	211	US-08-133-011-16	Sequence 16, Appl
146	5	10.9	97	3	US-08-816-977-23	Sequence 23, Appl	219	211	US-08-322-730A-16	Sequence 16, Appl
147	5	10.9	97	3	US-08-431-184-10	Sequence 10, Appl	220	211	US-08-387-874-16	Sequence 16, Appl
148	5	10.9	97	4	US-08-856-207A-431	Sequence 431, App	221	211	US-08-899-575-34	Sequence 34, Appl
149	5	10.9	107	1	US-08-369-796-16	Sequence 16, App	222	211	US-08-899-575-34	Sequence 34, Appl
150	5	10.9	107	2	US-08-852-091-16	Sequence 16, App	223	211	US-08-383-619-16	Sequence 16, Appl
151	5	10.9	107	5	PCT-US95-17025-36	Sequence 16, App	224	211	US-08-907-739-16	Sequence 16, Appl
152	5	10.9	109	1	US-07-662-193-5	Sequence 5, Appl	225	211	PCT-US93-08364-16	Sequence 16, Appl
153	5	10.9	109	1	US-07-807-529A-6	Sequence 6, Appl	226	212	PCT-US95-08743-34	Sequence 34, Appl
154	5	10.9	109	3	US-08-300-928C-8	Sequence 8, Appl	227	212	US-08-737-129A-2	Sequence 2, Appl
155	5	10.9	109	3	US-08-430-944D-8	Sequence 8, Appl	228	213	US-08-486-099-110	Sequence 110, App
156	5	10.9	109	3	US-08-430-014-8	Sequence 8, Appl	229	213	US-08-360-107A-120	Sequence 120, App
157	5	10.9	109	3	US-08-431-184-8	Sequence 8, Appl	230	213	US-08-484-223B-110	Sequence 110, App
158	5	10.9	109	5	PCT-US93-02462-6	Sequence 6, Appl	231	213	US-08-919-597-110	Sequence 110, App
159	5	10.9	111	1	US-07-662-193-4	Sequence 4, Appl	232	213	US-08-475-668A-110	Sequence 110, App
160	5	10.9	111	3	US-08-300-928C-6	Sequence 6, Appl	233	213	US-08-485-551A-110	Sequence 110, App
161	5	10.9	111	3	US-08-430-944D-6	Sequence 6, Appl	234	213	US-08-471-913A-110	Sequence 110, App
162	5	10.9	111	3	US-08-430-014-6	Sequence 6, Appl	235	213	US-08-485-264A-110	Sequence 110, App
163	5	10.9	111	3	US-08-431-184-6	Sequence 6, Appl	236	213	US-08-474-349A-110	Sequence 110, App
164	5	10.9	116	1	US-08-139-862-4	Sequence 4, Appl	237	214	US-08-277-231A-14	Sequence 14, Appl
165	5	10.9	117	4	US-09-046-479-2	Sequence 2, Appl	238	214	US-08-473-750-3	Sequence 3, Appl
166	5	10.9	120	2	US-08-446-345-27	Sequence 27, Appl	239	214	US-08-477-326-3	Sequence 3, Appl
167	5	10.9	123	1	US-08-082-623-3	Sequence 3, Appl	240	216	US-08-277-231A-13	Sequence 13, Appl
168	5	10.9	124	4	US-08-983-607-47	Sequence 47, Appl	241	216	US-08-473-750-2	Sequence 2, Appl
169	5	10.9	124	4	US-08-789-333F-41	Sequence 41, Appl	242	216	US-08-477-326-2	Sequence 2, Appl
170	5	10.9	125	4	US-08-751-359-13	Sequence 13, Appl	243	217	US-08-277-231A-2	Sequence 2, Appl
171	5	10.9	125	4	US-08-907-146-13	Sequence 13, Appl	244	217	US-08-473-750-5	Sequence 5, Appl
172	5	10.9	126	4	US-08-751-359-12	Sequence 12, Appl	245	217	US-08-477-326-5	Sequence 5, Appl
173	5	10.9	126	4	US-08-907-146-12	Sequence 12, Appl	246	217	PCT-US94-14106-59	Sequence 59, Appl

247	5	10.9	218	4	US-09-227-357-368	Sequence 368, Appl	320	5	10.9	273	4	US-09-352-302-4	Sequence 4, Appl
248	5	10.9	219	1	US-08-152-019A-31	Sequence 31, Appl	321	5	10.9	274	3	US-08-491-954-4	Sequence 4, Appl
249	5	10.9	219	1	US-08-132-019A-32	Sequence 32, Appl	322	5	10.9	276	2	US-08-701-935-1	Sequence 1, Appl
250	5	10.9	219	1	US-08-266-451B-24	Sequence 24, Appl	323	5	10.9	276	3	US-09-134-591-1	Sequence 1, Appl
251	5	10.9	219	2	US-08-460-309-18	Sequence 18, Appl	324	5	10.9	278	3	US-08-491-954-2	Sequence 2, Appl
252	5	10.9	219	2	US-08-748-725-24	Sequence 24, Appl	325	5	10.9	280	1	US-08-434-255-8	Sequence 8, Appl
253	5	10.9	219	2	US-08-125-077-18	Sequence 18, Appl	326	5	10.9	280	1	US-08-459-967-8	Sequence 8, Appl
254	5	10.9	220	4	US-09-052-089A-3	Sequence 3, Appl	327	5	10.9	280	1	US-08-460-327-8	Sequence 8, Appl
255	5	10.9	220	4	US-08-737-129A-6	Sequence 6, Appl	328	5	10.9	280	1	US-08-459-871-8	Sequence 8, Appl
256	5	10.9	222	4	US-09-181-183-28	Sequence 28, Appl	329	5	10.9	280	4	US-09-024-532-2	Sequence 2, Appl
257	5	10.9	222	4	US-09-277-700-28	Sequence 28, Appl	330	5	10.9	280	4	US-09-104-623A-2	Sequence 2, Appl
258	5	10.9	223	4	US-09-129-030-30	Sequence 30, Appl	331	5	10.9	282	5	PCT-US94-09752-3	Sequence 3, Appl
259	5	10.9	223	5	PCT-US94-14106-51	Sequence 51, Appl	332	5	10.9	284	3	US-08-491-954-3	Sequence 3, Appl
260	5	10.9	223	5	PCT-US94-14106-55	Sequence 55, Appl	333	5	10.9	287	3	US-08-457-245-3	Sequence 3, Appl
261	5	10.9	227	4	US-08-213-419B-13	Sequence 13, Appl	334	5	10.9	287	3	US-09-173-581-2	Sequence 2, Appl
262	5	10.9	228	4	US-09-181-958-2	Sequence 2, Appl	335	5	10.9	287	3	US-08-862-124-17	Sequence 17, Appl
263	5	10.9	229	2	US-08-121-436A-4	Sequence 4, Appl	336	5	10.9	287	4	US-09-420-915-2	Sequence 2, Appl
264	5	10.9	230	2	US-08-808-550-34	Sequence 34, Appl	337	5	10.9	288	3	US-09-100-804-14	Sequence 14, Appl
265	5	10.9	232	2	US-08-829-876-103	Sequence 103, Appl	338	5	10.9	292	3	US-08-851-190-3	Sequence 3, Appl
266	5	10.9	233	2	US-08-829-876-101	Sequence 101, Appl	339	5	10.9	293	4	US-08-438-745-4	Sequence 4, Appl
267	5	10.9	233	2	US-08-829-876-105	Sequence 105, Appl	340	5	10.9	293	4	US-08-438-745-6	Sequence 6, Appl
268	5	10.9	233	4	US-08-438-745-15	Sequence 15, Appl	341	5	10.9	293	4	US-09-219-019-4	Sequence 4, Appl
269	5	10.9	233	4	US-08-438-745-17	Sequence 17, Appl	342	5	10.9	293	4	US-09-219-019-6	Sequence 6, Appl
270	5	10.9	233	4	US-09-219-019-15	Sequence 15, Appl	343	5	10.9	293	5	PCT-US94-05669A-4	Sequence 4, Appl
271	5	10.9	233	4	US-09-219-019-17	Sequence 17, Appl	344	5	10.9	293	5	PCT-US94-05669A-6	Sequence 6, Appl
272	5	10.9	233	5	PCT-US94-05669A-15	Sequence 15, Appl	345	5	10.9	296	2	US-08-836-854-4	Sequence 4, Appl
273	5	10.9	233	5	PCT-US94-05669A-17	Sequence 17, Appl	346	5	10.9	297	1	US-08-534-910B-7	Sequence 7, Appl
274	5	10.9	234	4	US-09-040-483-5	Sequence 5, Appl	347	5	10.9	297	1	US-08-534-910B-10	Sequence 10, Appl
275	5	10.9	235	4	US-08-438-745-13	Sequence 13, Appl	348	5	10.9	297	3	US-09-173-581-1	Sequence 1, Appl
276	5	10.9	235	4	US-09-219-019-13	Sequence 13, Appl	349	5	10.9	297	4	US-09-420-915-1	Sequence 1, Appl
277	5	10.9	235	5	PCT-US94-05669A-13	Sequence 13, Appl	350	5	10.9	299	3	US-08-986-765-2	Sequence 2, Appl
278	5	10.9	236	1	US-08-158-682A-4	Sequence 4, Appl	351	5	10.9	299	4	US-08-935-307-2	Sequence 2, Appl
279	5	10.9	236	1	US-08-687-895-4	Sequence 4, Appl	352	5	10.9	299	4	US-09-524-641-2	Sequence 2, Appl
280	5	10.9	236	1	US-08-816-241-4	Sequence 4, Appl	353	5	10.9	299	4	US-08-965-130-2	Sequence 2, Appl
281	5	10.9	236	3	US-09-040-482-4	Sequence 4, Appl	354	5	10.9	300	1	US-08-148-910-1	Sequence 1, Appl
282	5	10.9	236	3	US-09-128-395-4	Sequence 5, Appl	355	5	10.9	300	1	US-08-448-937A-1	Sequence 1, Appl
283	5	10.9	239	2	US-08-916-902A-1	Sequence 1, Appl	356	5	10.9	300	4	US-08-158-735A-16	Sequence 16, Appl
284	5	10.9	239	2	US-09-213-389-1	Sequence 8, Appl	357	5	10.9	300	4	US-08-158-735A-17	Sequence 17, Appl
285	5	10.9	240	4	US-09-194-905-8	Sequence 8, Appl	358	5	10.9	301	3	US-08-169-318-142	Sequence 142, Appl
286	5	10.9	243	3	US-09-286-690-10	Sequence 10, Appl	359	5	10.9	301	3	US-08-468-609A-142	Sequence 142, Appl
287	5	10.9	246	2	US-08-553-497A-24	Sequence 24, Appl	360	5	10.9	301	5	PCT-US95-01185-142	Sequence 142, Appl
288	5	10.9	247	3	US-09-100-804-13	Sequence 13, Appl	361	5	10.9	304	4	US-08-862-124-14	Sequence 14, Appl
289	5	10.9	247	3	PCT-US94-07659-2	Sequence 2, Appl	362	5	10.9	310	1	US-08-363-255-6	Sequence 6, Appl
290	5	10.9	248	4	US-08-944-483-71	Sequence 71, Appl	363	5	10.9	312	1	US-08-247-908A-2	Sequence 2, Appl
291	5	10.9	251	2	US-08-933-750C-44	Sequence 44, Appl	364	5	10.9	312	1	US-08-453-942-2	Sequence 2, Appl
292	5	10.9	251	4	US-09-234-613-44	Sequence 44, Appl	365	5	10.9	312	5	US-08-926-885A-2	Sequence 2, Appl
293	5	10.9	251	4	US-08-944-483-28	Sequence 28, Appl	366	5	10.9	312	5	PCT-US94-05290-2	Sequence 2, Appl
294	5	10.9	254	2	US-08-792-824-4	Sequence 4, Appl	367	5	10.9	320	2	US-08-579-940-8	Sequence 8, Appl
295	5	10.9	254	2	US-08-792-824-7	Sequence 7, Appl	368	5	10.9	320	3	US-08-364-308-2	Sequence 2, Appl
296	5	10.9	254	2	US-08-792-824-10	Sequence 10, Appl	369	5	10.9	320	3	US-08-964-313-2	Sequence 2, Appl
297	5	10.9	254	2	US-08-792-824-13	Sequence 13, Appl	370	5	10.9	320	4	US-09-069-138-2	Sequence 2, Appl
298	5	10.9	261	4	US-08-927-219-55	Sequence 55, Appl	371	5	10.9	322	3	US-08-964-308-8	Sequence 8, Appl
299	5	10.9	264	4	US-08-904-284-3	Sequence 3, Appl	372	5	10.9	322	3	US-08-964-313-8	Sequence 8, Appl
300	5	10.9	268	1	US-08-431-387-4	Sequence 4, Appl	373	5	10.9	322	4	US-09-069-138-8	Sequence 8, Appl
301	5	10.9	268	2	US-07-857-224B-33	Sequence 33, Appl	374	5	10.9	324	2	US-08-579-940-7	Sequence 7, Appl
302	5	10.9	269	1	US-08-203-716-3	Sequence 3, Appl	375	5	10.9	325	4	US-08-878-474-9	Sequence 9, Appl
303	5	10.9	269	1	US-08-440-179-3	Sequence 3, Appl	376	5	10.9	328	1	US-08-229-287-4	Sequence 4, Appl
304	5	10.9	269	2	US-08-432-693-11	Sequence 11, Appl	377	5	10.9	328	2	US-08-977-767-1	Sequence 1, Appl
305	5	10.9	269	2	US-08-599-895-11	Sequence 11, Appl	378	5	10.9	331	1	US-08-094-533B-10	Sequence 10, Appl
306	5	10.9	269	2	US-08-483-806-2	Sequence 2, Appl	379	5	10.9	331	1	US-08-276-860A-10	Sequence 10, Appl
307	5	10.9	269	3	US-09-211-290-11	Sequence 11, Appl	380	5	10.9	331	1	US-08-444-393-10	Sequence 10, Appl
308	5	10.9	269	3	US-09-320-613-11	Sequence 11, Appl	381	5	10.9	331	2	US-08-799-913-10	Sequence 10, Appl
309	5	10.9	269	3	US-09-322-676-11	Sequence 11, Appl	382	5	10.9	331	2	US-08-711-893-10	Sequence 10, Appl
310	5	10.9	269	4	US-08-039-657-3	Sequence 3, Appl	383	5	10.9	331	2	US-09-150-200-10	Sequence 10, Appl
311	5	10.9	269	4	US-08-748-547-4	Sequence 4, Appl	384	5	10.9	331	3	US-09-150-201-10	Sequence 10, Appl
312	5	10.9	269	4	US-09-466-036A-11	Sequence 11, Appl	385	5	10.9	331	4	US-08-961-083-212	Sequence 212, Appl
313	5	10.9	269	4	US-09-451-905-11	Sequence 11, Appl	386	5	10.9	331	4	US-09-452-370-10	Sequence 10, Appl
314	5	10.9	269	5	PCT-US91-02339-1	Sequence 1, Appl	387	5	10.9	331	4	US-09-461-649-10	Sequence 10, Appl
315	5	10.9	270	1	US-08-347-471-4	Sequence 4, Appl	388	5	10.9	331	5	PCT-US94-08119-10	Sequence 10, Appl
316	5	10.9	270	2	US-09-035-095-4	Sequence 4, Appl	389	5	10.9	331	5	PCT-US94-08120-10	Sequence 10, Appl
317	5	10.9	270	2	US-08-809-494A-2	Sequence 2, Appl	390	5	10.9	331	5	PCT-US94-12913A-10	Sequence 10, Appl
318	5	10.9	270	4	US-09-352-302-2	Sequence 2, Appl	391	5	10.9	331	5	PCT-US95-15353-39	Sequence 39, Appl
319	5	10.9	273	2	US-08-809-494A-4	Sequence 4, Appl	392	5	10.9	331	5	PCT-US95-15353-41	Sequence 41, Appl

393	5	10.9	335	3	US-08-469-318-143	Sequence 143, App	466	5	10.9	397	1	US-08-459-871-4	Sequence 4, Appli
394	5	10.9	335	3	US-08-468-609A-143	Sequence 143, App	467	5	10.9	398	4	US-08-810-712-4	Sequence 4, Appli
395	5	10.9	335	5	PCT-US95-01185-143	Sequence 143, App	468	5	10.9	399	1	US-08-530-950-10	Sequence 10, Appl
396	5	10.9	337	3	US-08-469-318-148	Sequence 148, App	469	5	10.9	399	2	US-08-874-186-92	Sequence 92, Appl
397	5	10.9	337	3	US-08-468-609A-148	Sequence 148, App	470	5	10.9	399	4	US-08-888-429A-10	Sequence 10, Appl
398	5	10.9	337	3	US-08-480-070C-9	Sequence 9, Appli	471	5	10.9	399	4	US-09-149-879-10	Sequence 2, Appli
399	5	10.9	337	3	US-08-829-525-9	Sequence 9, Appli	472	5	10.9	399	4	US-09-553-498-2	Sequence 7, Appli
400	5	10.9	337	4	US-08-609-583A-9	Sequence 9, Appli	473	5	10.9	402	1	US-08-315-461-7	Sequence 4, Appli
401	5	10.9	337	4	US-08-937-399-9	Sequence 9, Appli	474	5	10.9	402	3	US-08-948-997-4	Sequence 2, Appli
402	5	10.9	337	5	PCT-US95-01185-148	Sequence 148, App	475	5	10.9	402	3	US-08-840-204-2	Sequence 4, Appli
403	5	10.9	342	3	US-08-978-456-2	Sequence 2, Appli	476	5	10.9	402	4	US-09-348-817A-4	Sequence 3, Appli
404	5	10.9	342	4	US-09-369-700-2	Sequence 2, Appli	477	5	10.9	402	4	US-09-026-408-3	Sequence 6, Appli
405	5	10.9	343	1	US-08-463-090B-11	Sequence 11, Appl	478	5	10.9	409	1	US-09-640-305-6	Sequence 2, Appli
406	5	10.9	344	4	US-08-961-083-132	Sequence 132, App	479	5	10.9	409	1	US-08-360-673-6	Sequence 6, Appli
407	5	10.9	345	2	US-08-446-345-40	Sequence 40, Appl	480	5	10.9	415	4	US-09-100-193-2	Sequence 2, Appli
408	5	10.9	349	3	US-08-469-318-139	Sequence 139, App	481	5	10.9	415	4	US-09-025-769B-280	Sequence 280, App
409	5	10.9	349	3	US-08-469-318-151	Sequence 151, App	482	5	10.9	416	1	US-08-073-807A-2	Sequence 2, Appli
410	5	10.9	349	3	US-08-468-609A-139	Sequence 139, App	483	5	10.9	416	3	US-08-858-876A-4	Sequence 4, Appli
411	5	10.9	349	3	US-08-468-609A-151	Sequence 151, App	484	5	10.9	416	4	US-09-472-880-4	Sequence 4, Appli
412	5	10.9	349	5	PCT-US95-01185-139	Sequence 139, App	485	5	10.9	417	3	US-08-705-771-18	Sequence 18, Appl
413	5	10.9	349	5	PCT-US95-01185-151	Sequence 151, App	486	5	10.9	424	1	US-08-247-908A-11	Sequence 11, Appl
414	5	10.9	353	1	US-08-073-807A-17	Sequence 17, Appl	487	5	10.9	424	1	US-08-453-942-11	Sequence 11, Appl
415	5	10.9	360	1	US-08-597-236-11	Sequence 11, Appl	488	5	10.9	424	2	US-08-926-885A-11	Sequence 11, Appl
416	5	10.9	360	1	US-08-746-682A-11	Sequence 11, Appl	489	5	10.9	424	5	PCT-US94-05290-11	Sequence 11, Appl
417	5	10.9	360	4	US-09-364-230-4	Sequence 4, Appli	490	5	10.9	428	3	US-09-118-319-5	Sequence 5, Appli
418	5	10.9	361	1	US-08-537-434-1	Sequence 1, Appli	491	5	10.9	430	1	US-08-785-076-3	Sequence 3, Appli
419	5	10.9	361	4	US-08-096-181A-8	Sequence 8, Appli	492	5	10.9	432	1	US-08-785-076-2	Sequence 2, Appli
420	5	10.9	361	5	PCT-US94-08326-8	Sequence 8, Appli	493	5	10.9	432	2	US-08-644-034A-2	Sequence 2, Appli
421	5	10.9	362	2	US-09-445-090-2	Sequence 2, Appli	494	5	10.9	434	2	US-08-844-134-2	Sequence 2, Appli
422	5	10.9	362	4	US-09-286-691-14	Sequence 14, Appl	495	5	10.9	434	3	US-09-128-192A-2	Sequence 2, Appli
423	5	10.9	362	4	US-09-687-147-14	Sequence 14, Appl	496	5	10.9	436	4	US-08-716-873-5	Sequence 5, Appli
424	5	10.9	363	1	US-08-484-105-20	Sequence 20, Appl	497	5	10.9	436	4	US-09-368-431-5	Sequence 5, Appli
425	5	10.9	363	1	US-08-484-106-20	Sequence 20, Appl	498	5	10.9	437	1	US-08-764-343-1	Sequence 1, Appli
426	5	10.9	363	2	US-08-646-590B-36	Sequence 36, Appl	499	5	10.9	437	1	US-08-810-116-10	Sequence 10, Appl
427	5	10.9	363	4	US-09-412-184-36	Sequence 36, Appl	500	5	10.9	437	2	US-07-930-548A-10	Sequence 10, Appl
428	5	10.9	367	4	US-08-213-419B-15	Sequence 15, Appl	501	5	10.9	437	2	US-08-989-925-3	Sequence 3, Appli
429	5	10.9	369	1	US-08-700-339-21	Sequence 21, Appl	502	5	10.9	438	2	US-08-677-049-9	Sequence 9, Appli
430	5	10.9	370	1	US-08-434-255-6	Sequence 6, Appli	503	5	10.9	438	4	US-09-167-239-5	Sequence 5, Appli
431	5	10.9	370	1	US-08-459-967-6	Sequence 6, Appli	504	5	10.9	439	4	US-08-716-873-2	Sequence 2, Appli
432	5	10.9	370	1	US-08-460-327-6	Sequence 6, Appli	505	5	10.9	439	4	US-08-716-873-4	Sequence 4, Appli
433	5	10.9	370	1	US-08-459-871-6	Sequence 6, Appli	506	5	10.9	439	4	US-09-368-431-2	Sequence 2, Appli
434	5	10.9	370	3	US-08-911-321-10	Sequence 10, Appl	507	5	10.9	439	4	US-09-368-431-4	Sequence 4, Appli
435	5	10.9	371	2	US-08-948-176-26	Sequence 26, Appl	508	5	10.9	442	1	US-08-363-255-4	Sequence 4, Appli
436	5	10.9	374	2	US-08-928-692-51	Sequence 51, Appl	509	5	10.9	442	1	US-08-363-255-11	Sequence 11, Appl
437	5	10.9	374	3	US-08-609-236-6	Sequence 6, Appli	510	5	10.9	444	4	US-09-253-292C-27	Sequence 27, Appl
438	5	10.9	375	2	US-08-494-151-14	Sequence 14, Appl	511	5	10.9	445	1	US-08-353-400-33	Sequence 33, Appl
439	5	10.9	375	3	US-09-106-217-16	Sequence 16, Appl	512	5	10.9	445	1	US-08-363-255-5	Sequence 5, Appli
440	5	10.9	375	4	US-09-171-337A-7	Sequence 7, Appli	513	5	10.9	445	1	US-08-363-255-12	Sequence 12, Appli
441	5	10.9	375	1	US-08-171-337A-8	Sequence 8, Appli	514	5	10.9	447	2	US-08-902-585-2	Sequence 2, Appli
442	5	10.9	376	1	US-08-588-113-2	Sequence 2, Appli	515	5	10.9	447	6	5455030-1	Patent No. 5455030
443	5	10.9	377	3	US-09-106-217-2	Sequence 2, Appli	516	5	10.9	448	1	US-08-570-157-3	Sequence 3, Appli
444	5	10.9	379	3	US-08-840-204-3	Sequence 3, Appli	517	5	10.9	449	3	US-08-987-743-7	Sequence 7, Appli
445	5	10.9	379	4	US-09-193-266-1	Sequence 1, Appli	518	5	10.9	450	4	US-09-268-992-41	Sequence 41, Appl
446	5	10.9	380	1	US-08-073-807A-18	Sequence 18, Appl	519	5	10.9	451	4	US-09-357-251-35	Sequence 35, Appl
447	5	10.9	382	4	US-08-213-419B-19	Sequence 19, Appl	520	5	10.9	452	1	US-08-290-978A-5	Sequence 5, Appli
448	5	10.9	385	1	US-08-036-210-5	Sequence 5, Appli	521	5	10.9	452	2	US-08-780-869-5	Sequence 2, Appli
449	5	10.9	385	2	US-08-449-609-5	Sequence 5, Appli	522	5	10.9	459	4	US-09-491-785-2	Sequence 2, Appli
450	5	10.9	390	1	US-08-121-714-6	Sequence 6, Appli	523	5	10.9	461	2	US-08-463-587A-26	Sequence 26, Appl
451	5	10.9	390	1	US-08-477-108A-6	Sequence 6, Appli	524	5	10.9	461	2	US-08-463-667A-4	Sequence 4, Appli
452	5	10.9	390	2	US-08-477-113-6	Sequence 6, Appli	525	5	10.9	461	3	US-08-923-854-26	Sequence 26, Appl
453	5	10.9	390	5	PCT-US93-08322-6	Sequence 6, Appli	526	5	10.9	461	5	PCT-US91-09133-27	Sequence 27, Appl
454	5	10.9	392	2	US-08-799-173A-7	Sequence 7, Appli	527	5	10.9	464	1	US-08-353-400-36	Sequence 36, Appl
455	5	10.9	392	2	US-08-886-152-1	Sequence 1, Appli	528	5	10.9	464	2	US-08-553-619B-9	Sequence 9, Appli
456	5	10.9	392	2	US-08-886-152-3	Sequence 3, Appli	529	5	10.9	464	3	US-09-056-783-2	Sequence 2, Appli
457	5	10.9	394	2	US-08-530-290-23	Sequence 23, Appl	530	5	10.9	466	4	US-09-268-992-39	Sequence 39, Appl
458	5	10.9	394	4	US-09-144-914-4	Sequence 4, Appli	531	5	10.9	467	4	US-09-086-483A-6	Sequence 6, Appli
459	5	10.9	397	1	US-08-434-255-2	Sequence 2, Appli	532	5	10.9	467	4	US-08-495-484-12	Sequence 12, Appl
460	5	10.9	397	1	US-08-434-255-4	Sequence 4, Appli	533	5	10.9	468	2	US-08-303-569B-7	Sequence 7, Appli
461	5	10.9	397	1	US-08-459-967-2	Sequence 2, Appli	534	5	10.9	468	2	US-08-116-247-7	Sequence 7, Appli
462	5	10.9	397	1	US-08-459-967-4	Sequence 4, Appli	535	5	10.9	468	2	US-09-013-855A-2	Sequence 2, Appli
463	5	10.9	397	1	US-08-460-327-2	Sequence 2, Appli	536	5	10.9	469	2	US-08-968-751-2	Sequence 2, Appli
464	5	10.9	397	1	US-08-460-327-4	Sequence 4, Appli	537	5	10.9	469	4	US-09-052-089A-1	Sequence 1, Appli
465	5	10.9	397	1	US-08-459-871-2	Sequence 2, Appli	538	5	10.9	472	1	US-08-749-903-1	Sequence 1, Appli

539	5	10.9	472	1	US-08-749-903-3	Sequence 3, Appli	612	5	10.9	542	4	US-08-969-683A-16	Sequence 16, Appli
540	5	10.9	472	4	US-09-088-641-1	Sequence 1, Appli	613	5	10.9	542	4	US-09-297-928-12	Sequence 12, Appli
541	5	10.9	472	4	US-09-088-641-3	Sequence 3, Appli	614	5	10.9	543	2	US-08-469-412A-7	Sequence 7, Appli
542	5	10.9	473	4	US-08-857-076-99	Sequence 99, Appl	615	5	10.9	543	4	US-08-697-610-2	Sequence 2, Appli
543	5	10.9	473	4	US-09-171-945-131	Sequence 131, App	616	5	10.9	543	4	US-09-021-715-7	Sequence 7, Appli
544	5	10.9	477	2	US-08-432-016-3	Sequence 3, Appli	617	5	10.9	543	4	US-08-349-357-2	Sequence 2, Appli
545	5	10.9	477	2	US-08-684-594-3	Sequence 3, Appli	618	5	10.9	545	1	US-08-133-347-2	Sequence 2, Appli
546	5	10.9	479	1	US-08-484-105-10	Sequence 10, Appl	619	5	10.9	549	2	US-08-500-635A-12	Sequence 12, Appli
547	5	10.9	479	1	US-08-484-106-10	Sequence 10, Appl	620	5	10.9	549	2	US-09-167-151-12	Sequence 12, Appli
548	5	10.9	479	4	US-09-126-109-8	Sequence 8, Appli	621	5	10.9	550	1	US-08-143-219-25	Sequence 25, Appli
549	5	10.9	480	2	US-08-828-488-8	Sequence 8, Appli	622	5	10.9	551	2	US-08-436-771-9	Sequence 9, Appli
550	5	10.9	480	3	US-09-189-035-5	Sequence 5, Appli	623	5	10.9	551	2	US-08-434-998-9	Sequence 9, Appli
551	5	10.9	480	4	US-09-382-086-5	Sequence 5, Appli	624	5	10.9	551	2	US-08-487-797-9	Sequence 9, Appli
552	5	10.9	482	1	US-08-382-160-5	Sequence 5, Appli	625	5	10.9	551	2	US-09-067-351-2	Sequence 2, Appli
553	5	10.9	482	3	US-09-135-639-2	Sequence 2, Appli	626	5	10.9	551	4	US-09-360-490-2	Sequence 2, Appli
554	5	10.9	484	1	US-08-358-150-7	Sequence 7, Appli	627	5	10.9	551	5	PCT-US95-020581-9	Sequence 9, Appli
555	5	10.9	492	3	US-08-724-466B-2	Sequence 2, Appli	628	5	10.9	558	1	US-08-285-440-6	Sequence 6, Appli
556	5	10.9	492	4	US-08-882-164D-2	Sequence 2, Appli	629	5	10.9	558	1	US-08-630-349-6	Sequence 6, Appli
557	5	10.9	493	1	US-07-615-448A-7	Sequence 7, Appli	630	5	10.9	563	4	US-09-311-924-2	Sequence 2, Appli
558	5	10.9	493	1	US-08-196-361-7	Sequence 7, Appli	631	5	10.9	567	4	US-08-927-219-131	Sequence 131, App
559	5	10.9	493	2	US-08-446-934-7	Sequence 7, Appli	632	5	10.9	574	2	US-08-836-854-21	Sequence 21, Appli
560	5	10.9	493	2	US-08-448-128-7	Sequence 7, Appli	633	5	10.9	574	4	US-09-463-296-1	Sequence 1, Appli
561	5	10.9	493	3	US-08-948-703-7	Sequence 7, Appli	634	5	10.9	577	1	US-08-484-105-24	Sequence 24, Appli
562	5	10.9	497	2	US-09-047-125-3	Sequence 3, Appli	635	5	10.9	577	1	US-08-484-106-24	Sequence 24, Appli
563	5	10.9	497	3	US-07-736-335E-3	Sequence 3, Appli	636	5	10.9	580	1	US-08-309-512-6	Sequence 6, Appli
564	5	10.9	498	4	US-09-058-260-16	Sequence 16, Appl	637	5	10.9	580	2	US-08-591-079-2	Sequence 2, Appli
565	5	10.9	501	1	US-08-149-105-15	Sequence 15, Appl	638	5	10.9	580	5	PCT-US92-08756A-6	Sequence 6, Appli
566	5	10.9	501	1	US-08-149-105-15	Sequence 15, Appl	639	5	10.9	581	3	US-08-619-812-6	Sequence 6, Appli
567	5	10.9	501	1	US-08-317-847-15	Sequence 15, Appl	640	5	10.9	584	1	US-08-261-822A-8	Sequence 8, Appli
568	5	10.9	501	1	US-08-317-847-17	Sequence 17, Appl	641	5	10.9	584	5	PCT-US95-07744A-8	Sequence 8, Appli
569	5	10.9	503	2	US-08-781-802-2	Sequence 2, Appli	642	5	10.9	585	1	US-08-477-674-10	Sequence 10, Appli
570	5	10.9	503	2	US-08-781-802-10	Sequence 10, Appl	643	5	10.9	585	1	US-08-791-10	Sequence 10, Appli
571	5	10.9	503	2	US-08-781-802-12	Sequence 12, Appl	644	5	10.9	585	2	US-08-316-714-10	Sequence 10, Appli
572	5	10.9	503	4	US-09-382-256-16	Sequence 16, Appl	645	5	10.9	585	3	US-08-473-673-10	Sequence 10, Appli
573	5	10.9	503	4	US-08-694-078-2	Sequence 2, Appli	646	5	10.9	599	1	US-08-442-542-18	Sequence 18, Appli
574	5	10.9	503	4	US-09-058-260-2	Sequence 2, Appli	647	5	10.9	599	3	US-08-765-469-18	Sequence 18, Appli
575	5	10.9	503	4	US-09-058-260-10	Sequence 10, Appl	648	5	10.9	600	6	5240706-1	Patent No. 5240706
576	5	10.9	503	4	US-09-058-260-12	Sequence 12, Appl	649	5	10.9	612	2	US-08-746-283-31	Sequence 31, Appli
577	5	10.9	503	4	US-09-058-260-28	Sequence 28, Appl	650	5	10.9	612	2	US-08-746-257A-29	Sequence 29, Appli
578	5	10.9	503	4	US-09-058-260-30	Sequence 30, Appl	651	5	10.9	613	2	US-08-746-283-1	Sequence 1, Appli
579	5	10.9	503	4	US-09-395-115-10	Sequence 10, Appl	652	5	10.9	613	2	US-08-746-283-3	Sequence 3, Appli
580	5	10.9	503	4	US-08-123-934A-10	Sequence 10, Appl	653	5	10.9	613	2	US-08-746-283-5	Sequence 5, Appli
581	5	10.9	503	4	US-08-436-265-10	Sequence 10, Appl	654	5	10.9	613	2	US-08-746-257A-1	Sequence 1, Appli
582	5	10.9	503	5	PCT-US94-10080-10	Sequence 10, Appl	655	5	10.9	615	3	US-08-860-091A-4	Sequence 4, Appli
583	5	10.9	505	4	US-09-382-256-8	Sequence 8, Appli	656	5	10.9	623	4	US-09-104-068-4	Sequence 4, Appli
584	5	10.9	505	4	US-09-382-256-16	Sequence 16, Appl	657	5	10.9	626	2	US-08-596-300A-7	Sequence 7, Appli
585	5	10.9	505	4	US-09-395-115-8	Sequence 8, Appli	658	5	10.9	626	2	US-08-596-300A-14	Sequence 14, Appli
586	5	10.9	505	4	US-09-395-115-16	Sequence 16, Appl	659	5	10.9	630	4	US-09-704-449-2	Sequence 2, Appli
587	5	10.9	505	4	US-08-123-934A-8	Sequence 8, Appli	660	5	10.9	637	4	US-09-104-068-21	Sequence 21, Appli
588	5	10.9	505	4	US-08-436-265-8	Sequence 8, Appli	661	5	10.9	638	4	US-09-228-986-74	Sequence 74, Appli
589	5	10.9	505	4	US-08-436-265-16	Sequence 16, Appl	662	5	10.9	639	1	US-08-466-390-2	Sequence 2, Appli
590	5	10.9	505	5	PCT-US94-10080-8	Sequence 8, Appli	663	5	10.9	639	1	US-08-470-950-2	Sequence 2, Appli
591	5	10.9	508	2	US-08-724-281-1	Sequence 1, Appli	664	5	10.9	639	1	US-08-467-781-2	Sequence 2, Appli
592	5	10.9	513	1	US-08-390-162-6	Sequence 6, Appli	665	5	10.9	639	1	US-08-195-487-2	Sequence 2, Appli
593	5	10.9	513	1	US-08-685-945B-6	Sequence 6, Appli	666	5	10.9	639	2	US-08-483-924-2	Sequence 2, Appli
594	5	10.9	513	3	US-08-948-564-2	Sequence 2, Appli	667	5	10.9	639	2	PCT-US93-06160-2	Sequence 2, Appli
595	5	10.9	516	4	US-08-927-219-140	Sequence 140, App	668	5	10.9	641	1	US-07-718-535-3	Sequence 3, Appli
596	5	10.9	516	4	US-09-201-641-6	Sequence 6, Appli	669	5	10.9	641	1	US-08-161-999-3	Sequence 3, Appli
597	5	10.9	524	1	US-08-624-125-2	Sequence 2, Appli	670	5	10.9	643	4	US-09-178-252-25	Sequence 25, Appli
598	5	10.9	524	1	US-08-624-125-21	Sequence 21, Appl	671	5	10.9	647	4	US-09-423-439-60	Sequence 60, Appl
599	5	10.9	525	4	US-08-764-870-7	Sequence 7, Appli	672	5	10.9	655	1	US-08-148-910-12	Sequence 12, Appli
600	5	10.9	525	4	US-08-980-115-7	Sequence 7, Appli	673	5	10.9	655	1	US-08-448-937A-12	Sequence 12, Appli
601	5	10.9	527	4	US-09-311-924-4	Sequence 4, Appli	674	5	10.9	670	4	US-08-980-080-2	Sequence 2, Appli
602	5	10.9	532	1	US-08-283-440-5	Sequence 5, Appli	675	5	10.9	685	2	US-08-678-989-1	Sequence 1, Appli
603	5	10.9	532	1	US-08-630-349-5	Sequence 5, Appli	676	5	10.9	685	3	US-09-136-282-2	Sequence 2, Appli
604	5	10.9	533	1	US-07-952-800-2	Sequence 2, Appli	677	5	10.9	685	4	US-09-272-796-1	Sequence 1, Appli
605	5	10.9	533	2	US-08-770-544-4	Sequence 4, Appli	678	5	10.9	685	4	US-09-505-744-2	Sequence 2, Appli
606	5	10.9	535	3	US-08-813-574-2	Sequence 2, Appli	679	5	10.9	686	2	US-08-768-301-4	Sequence 4, Appli
607	5	10.9	539	2	US-08-818-024-4	Sequence 4, Appli	680	5	10.9	694	2	US-08-895-522-3	Sequence 3, Appli
608	5	10.9	539	4	US-09-334-775A-4	Sequence 4, Appli	681	5	10.9	694	3	US-08-195-391-3	Sequence 3, Appli
609	5	10.9	539	4	US-08-789-275-6	Sequence 6, Appli	682	5	10.9	706	1	US-08-484-105-16	Sequence 16, Appli
610	5	10.9	541	4	US-09-627-376-11	Sequence 11, Appl	683	5	10.9	706	1	US-08-484-106-16	Sequence 16, Appli
611	5	10.9	542	3	US-08-968-563-16	Sequence 16, Appl	684	5	10.9	713	4	US-09-346-237-9	Sequence 9, Appli

685	5	10.9	719	4	US-09-588-256-2	Sequence 2, Appli	758	5	10.9	888	2	US-08-861-464-6	Sequence 6, Appli
686	5	10.9	720	2	US-08-840-236-1	Sequence 1, Appli	759	5	10.9	888	2	US-08-396-001-6	Sequence 6, Appli
687	5	10.9	720	1	US-08-505-448A-1	Sequence 1, Appli	760	5	10.9	888	2	US-09-323-433A-6	Sequence 6, Appli
688	5	10.9	725	2	US-08-813-940-25	Sequence 25, Appli	761	5	10.9	903	4	US-08-209-521-24	Sequence 24, Appli
689	5	10.9	726	3	US-09-129-075-4	Sequence 4, Appli	762	5	10.9	903	4	US-08-961-810-134	Sequence 134, App
690	5	10.9	726	4	US-09-346-237-3	Sequence 3, Appli	763	5	10.9	903	4	US-08-352-902D-134	Sequence 134, App
691	5	10.9	726	4	US-09-346-237-13	Sequence 13, Appli	764	5	10.9	903	4	US-09-228-986-78	Sequence 78, Appli
692	5	10.9	729	1	US-07-640-029-3	Sequence 3, Appli	765	5	10.9	907	2	US-09-010-928B-4	Sequence 4, Appli
693	5	10.9	731	1	US-07-921-807B-5	Sequence 5, Appli	766	5	10.9	908	3	US-08-699-103B-12	Sequence 12, Appli
694	5	10.9	731	1	US-08-441-944A-5	Sequence 5, Appli	767	5	10.9	908	4	US-09-229-059-12	Sequence 12, Appli
695	5	10.9	731	1	US-08-439-992A-3	Sequence 3, Appli	768	5	10.9	921	4	US-09-206-800-9	Sequence 9, Appli
696	5	10.9	733	1	US-07-840-029-4	Sequence 4, Appli	769	5	10.9	943	4	US-09-056-556-204	Sequence 204, App
697	5	10.9	733	1	US-07-921-807B-6	Sequence 6, Appli	770	5	10.9	948	1	US-08-698-551-14	Sequence 14, Appli
698	5	10.9	733	1	US-08-441-944A-6	Sequence 6, Appli	771	5	10.9	948	2	US-08-602-228-14	Sequence 14, Appli
699	5	10.9	733	4	US-08-439-992A-4	Sequence 4, Appli	772	5	10.9	948	2	US-08-533-901B-14	Sequence 14, Appli
700	5	10.9	734	1	US-08-276-099A-16	Sequence 16, Appli	773	5	10.9	948	2	US-08-839-032A-14	Sequence 14, Appli
701	5	10.9	734	1	US-08-781-890-16	Sequence 16, Appli	774	5	10.9	948	2	US-08-839-031A-14	Sequence 14, Appli
702	5	10.9	745	2	US-08-674-887A-8	Sequence 8, Appli	775	5	10.9	948	5	PCT-US95-12724-14	Sequence 3, Appli
703	5	10.9	745	2	US-09-010-928B-28	Sequence 28, Appli	776	5	10.9	956	2	US-08-897-443-3	Sequence 2, Appli
704	5	10.9	745	3	US-08-951-844-8	Sequence 8, Appli	777	5	10.9	966	1	US-08-571-758-2	Sequence 2, Appli
705	5	10.9	747	2	US-08-895-522-1	Sequence 1, Appli	778	5	10.9	966	1	US-08-909-984A-2	Sequence 2, Appli
706	5	10.9	747	3	US-09-195-391-1	Sequence 1, Appli	779	5	10.9	966	1	US-08-909-983-2	Sequence 2, Appli
707	5	10.9	756	1	US-08-434-730-16	Sequence 16, Appli	780	5	10.9	972	3	US-08-750-141A-2	Sequence 3, Appli
708	5	10.9	758	4	US-09-413-814-32	Sequence 32, Appli	781	5	10.9	984	1	US-08-257-073-3	Sequence 3, Appli
709	5	10.9	774	1	US-07-731-157A-7	Sequence 7, Appli	782	5	10.9	984	2	US-08-184-009-120	Sequence 120, App
710	5	10.9	774	2	US-08-541-780-7	Sequence 7, Appli	783	5	10.9	984	4	US-08-458-356-120	Sequence 120, App
711	5	10.9	786	3	US-09-012-710-11	Sequence 11, Appli	784	5	10.9	984	4	US-08-460-736-120	Sequence 120, App
712	5	10.9	786	4	US-09-556-273-11	Sequence 11, Appli	785	5	10.9	989	4	US-08-213-419B-2	Sequence 2, Appli
713	5	10.9	788	1	US-08-232-538-15	Sequence 15, Appli	786	5	10.9	989	4	US-08-213-419B-2	Sequence 2, Appli
714	5	10.9	788	2	US-08-786-164-15	Sequence 15, Appli	787	5	10.9	997	2	US-08-387-942C-4	Sequence 4, Appli
715	5	10.9	793	3	US-09-012-710-10	Sequence 10, Appli	788	5	10.9	1050	3	US-09-045-632-49	Sequence 49, Appli
716	5	10.9	793	3	US-09-556-273-10	Sequence 10, Appli	789	5	10.9	1050	3	US-09-045-632-50	Sequence 50, Appli
717	5	10.9	794	1	US-08-366-276-2	Sequence 2, Appli	790	5	10.9	1066	4	US-09-004-838-24	Sequence 24, Appli
718	5	10.9	794	1	US-08-393-333-2	Sequence 2, Appli	791	5	10.9	1115	3	US-08-323-477-2	Sequence 2, Appli
719	5	10.9	794	3	US-09-087-465-10	Sequence 10, Appli	792	5	10.9	1123	4	US-09-408-865-1	Sequence 1, Appli
720	5	10.9	799	3	US-08-909-954-4	Sequence 4, Appli	793	5	10.9	1127	4	US-09-150-460B-11	Sequence 11, Appli
721	5	10.9	807	1	US-07-862-021B-10	Sequence 10, Appli	794	5	10.9	1139	1	US-08-537-210A-4	Sequence 4, Appli
722	5	10.9	807	1	US-08-313-288B-10	Sequence 10, Appli	795	5	10.9	1139	1	US-09-113-825-4	Sequence 4, Appli
723	5	10.9	807	5	PCT-US93-03164-10	Sequence 10, Appli	796	5	10.9	1141	1	US-08-131-365B-54	Sequence 54, Appli
724	5	10.9	808	2	US-08-629-291A-33	Sequence 33, Appli	797	5	10.9	1141	2	US-08-668-123-54	Sequence 54, Appli
725	5	10.9	808	2	US-08-658-335B-33	Sequence 33, Appli	798	5	10.9	1146	4	US-08-914-999-6	Sequence 6, Appli
726	5	10.9	811	4	US-09-199-637A-93	Sequence 93, Appli	799	5	10.9	1147	1	US-08-131-365B-38	Sequence 38, Appli
727	5	10.9	814	1	US-08-233-788A-42	Sequence 42, Appli	800	5	10.9	1147	1	US-08-668-123-38	Sequence 38, Appli
728	5	10.9	817	1	US-07-640-029-2	Sequence 2, Appli	801	5	10.9	1164	2	US-08-589-756-1	Sequence 1, Appli
729	5	10.9	820	1	US-07-921-807B-3	Sequence 3, Appli	802	5	10.9	1164	4	US-09-206-800-1	Sequence 1, Appli
730	5	10.9	820	1	US-08-441-944A-3	Sequence 3, Appli	803	5	10.9	1164	4	US-09-206-898-1	Sequence 1, Appli
731	5	10.9	820	1	US-08-166-717D-6	Sequence 6, Appli	804	5	10.9	1186	4	US-09-178-252-23	Sequence 23, Appli
732	5	10.9	820	2	US-08-380-182-23	Sequence 23, Appli	805	5	10.9	1203	4	US-09-075-272-4	Sequence 4, Appli
733	5	10.9	820	4	US-08-439-992A-1	Sequence 1, Appli	806	5	10.9	1222	4	US-09-004-838-119	Sequence 119, App
734	5	10.9	822	1	US-07-997-133-1	Sequence 1, Appli	807	5	10.9	1227	1	US-08-448-170-8	Sequence 8, Appli
735	5	10.9	822	1	US-07-921-807B-4	Sequence 4, Appli	808	5	10.9	1227	4	US-08-961-803-9	Sequence 9, Appli
736	5	10.9	822	1	US-08-459-296-2	Sequence 2, Appli	809	5	10.9	1245	1	US-08-158-232-8	Sequence 8, Appli
737	5	10.9	822	1	US-08-441-944A-4	Sequence 4, Appli	810	5	10.9	1245	2	US-08-304-636-8	Sequence 8, Appli
738	5	10.9	822	2	US-08-451-822A-12	Sequence 12, Appli	811	5	10.9	1245	2	US-08-611-928-8	Sequence 8, Appli
739	5	10.9	822	4	US-08-439-992A-2	Sequence 2, Appli	812	5	10.9	1245	3	US-09-173-891-8	Sequence 8, Appli
740	5	10.9	822	1	US-08-323-430-12	Sequence 12, Appli	813	5	10.9	1248	3	US-08-726-214-16	Sequence 16, Appli
741	5	10.9	832	1	US-08-431-560-1	Sequence 1, Appli	814	5	10.9	1257	2	US-08-750-152A-2	Sequence 2, Appli
742	5	10.9	832	1	US-08-463-345-1	Sequence 1, Appli	815	5	10.9	1275	3	US-09-120-513-2	Sequence 2, Appli
743	5	10.9	835	4	US-09-291-835-1	Sequence 2, Appli	816	5	10.9	1275	3	US-09-450-105-2	Sequence 2, Appli
744	5	10.9	835	3	US-08-699-103B-10	Sequence 10, Appli	817	5	10.9	1279	2	US-08-784-649A-2	Sequence 2, Appli
745	5	10.9	853	4	US-09-229-059-10	Sequence 10, Appli	818	5	10.9	1280	2	US-08-583-276-19	Sequence 19, Appli
746	5	10.9	859	1	US-08-395-580-2	Sequence 2, Appli	819	5	10.9	1280	2	US-08-752-447-2	Sequence 2, Appli
747	5	10.9	859	5	PCT-US95-02792-2	Sequence 2, Appli	820	5	10.9	1280	6	5206352-4	Patent No. 5206352
748	5	10.9	870	1	US-08-785-241-4	Sequence 4, Appli	821	5	10.9	1290	1	US-08-138-641-2	Sequence 2, Appli
749	5	10.9	870	2	US-09-010-928B-2	Sequence 2, Appli	822	5	10.9	1290	1	US-08-138-133-2	Sequence 2, Appli
750	5	10.9	879	1	US-08-220-151-2	Sequence 2, Appli	823	5	10.9	1311	1	US-08-340-011-5	Sequence 5, Appli
751	5	10.9	879	1	US-08-220-151-3	Sequence 3, Appli	824	5	10.9	1311	3	US-08-901-710-5	Sequence 5, Appli
752	5	10.9	879	1	US-08-413-118-2	Sequence 2, Appli	825	5	10.9	1323	4	US-09-004-838-90	Sequence 90, Appli
753	5	10.9	879	1	US-08-413-118-3	Sequence 3, Appli	826	5	10.9	1337	3	US-08-854-585-2	Sequence 2, Appli
754	5	10.9	879	1	US-08-413-118-106	Sequence 106, App	827	5	10.9	1337	5	PCT-US95-05512-2	Sequence 2, Appli
755	5	10.9	879	3	US-08-473-446-2	Sequence 3, Appli	828	5	10.9	1338	1	US-08-750-141A-3	Sequence 3, Appli
756	5	10.9	879	3	US-08-473-446-3	Sequence 3, Appli	829	5	10.9	1356	1	US-08-810-116-8	Sequence 8, Appli
757	5	10.9	879	3	US-08-473-446-106	Sequence 106, App	830	5	10.9	1356	2	US-07-930-548A-8	Sequence 8, Appli

831	5	10.9	1356	4	US-09-098-707A-2	Sequence 2, Appl	904	5	10.9	2446	5	PCT-US93-12687-2	Sequence 2, Appl
832	5	10.9	1356	4	US-09-770-170-6	Sequence 6, Appl	905	5	10.9	2482	1	US-08-328-254-6	Sequence 6, Appl
833	5	10.9	1356	4	US-09-483-539-2	Sequence 2, Appl	906	5	10.9	2516	3	US-08-374-077C-2	Sequence 2, Appl
834	5	10.9	1362	2	US-08-874-678-33	Sequence 33, Appl	907	5	10.9	2516	2	US-08-895-590-2	Sequence 2, Appl
835	5	10.9	1362	3	US-08-643-839-33	Sequence 33, Appl	908	5	10.9	2627	2	US-08-751-189-3	Sequence 3, Appl
836	5	10.9	1363	1	US-08-425-061-23	Sequence 23, Appl	909	5	10.9	2627	2	US-09-060-836-3	Sequence 3, Appl
837	5	10.9	1363	2	US-08-825-886-23	Sequence 23, Appl	910	5	10.9	2627	2	US-09-184-445-3	Sequence 3, Appl
838	5	10.9	1426	4	US-09-136-574A-43	Sequence 43, Appl	911	5	10.9	2629	2	US-08-751-189-4	Sequence 4, Appl
839	5	10.9	1434	2	US-08-540-406-10	Sequence 10, Appl	912	5	10.9	2629	2	US-09-060-836-4	Sequence 4, Appl
840	5	10.9	1434	3	US-08-656-055-10	Sequence 10, Appl	913	5	10.9	2629	2	US-09-184-445-4	Sequence 4, Appl
841	5	10.9	1434	4	US-08-954-668-10	Sequence 10, Appl	914	5	10.9	2703	1	US-08-185-432-19	Sequence 19, Appl
842	5	10.9	1434	5	PCT-US95-13233-10	Sequence 10, Appl	915	5	10.9	2713	5	PCT-US96-01735-1	Sequence 1, Appl
843	5	10.9	1437	3	US-09-061-400-2	Sequence 2, Appl	916	5	10.9	2763	3	US-08-496-944-2	Sequence 2, Appl
844	5	10.9	1453	2	US-09-001-273-2	Sequence 2, Appl	917	5	10.9	2802	4	US-09-342-331-1	Sequence 1, Appl
845	5	10.9	1453	4	US-08-843-459A-2	Sequence 2, Appl	918	5	10.9	3072	4	US-09-413-814-93	Sequence 93, Appl
846	5	10.9	1463	1	US-08-220-603A-11	Sequence 11, Appl	919	5	10.9	3079	4	US-09-413-814-80	Sequence 80, Appl
847	5	10.9	1494	3	US-08-755-587-186	Sequence 186, Appl	920	5	10.9	3118	2	US-08-457-273B-8	Sequence 8, Appl
848	5	10.9	1572	2	US-08-290-731C-5	Sequence 5, Appl	921	5	10.9	3119	1	US-08-246-982A-16	Sequence 16, Appl
849	5	10.9	1581	4	US-09-110-517-2	Sequence 2, Appl	922	5	10.9	3119	1	US-08-453-265-16	Sequence 16, Appl
850	5	10.9	1596	3	US-09-356-952-3	Sequence 3, Appl	923	5	10.9	3144	1	US-08-246-982A-6	Sequence 6, Appl
851	5	10.9	1627	1	US-07-665-732E-9	Sequence 9, Appl	924	5	10.9	3144	1	US-08-453-265-6	Sequence 6, Appl
852	5	10.9	1786	2	US-08-477-451-16	Sequence 16, Appl	925	5	10.9	3144	2	US-08-457-273B-42	Sequence 42, Appl
853	5	10.9	1786	4	US-08-973-462-8	Sequence 8, Appl	926	5	10.9	3144	3	US-08-556-413-21	Sequence 21, Appl
854	5	10.9	1817	1	US-09-004-838-125	Sequence 125, Appl	927	5	10.9	3144	4	US-09-041-886-15	Sequence 15, Appl
855	5	10.9	1852	1	US-08-425-061-24	Sequence 24, Appl	928	5	10.9	3174	2	US-08-477-451-3	Sequence 3, Appl
856	5	10.9	1852	2	US-08-825-886-24	Sequence 24, Appl	929	5	10.9	3248	1	US-08-353-700-1	Sequence 1, Appl
857	5	10.9	1863	1	US-08-425-061-16	Sequence 16, Appl	930	5	10.9	3248	5	PCT-US95-16216-1	Sequence 1, Appl
858	5	10.9	1863	1	US-08-598-591-2	Sequence 2, Appl	931	5	10.9	7257	3	US-09-335-409-5	Sequence 5, Appl
859	5	10.9	1863	1	US-08-480-784-2	Sequence 2, Appl	932	5	10.9	7257	4	US-09-568-102-5	Sequence 5, Appl
860	5	10.9	1863	1	US-08-483-553-2	Sequence 2, Appl	933	5	10.9	7257	4	US-09-567-969-5	Sequence 5, Appl
861	5	10.9	1863	1	US-08-487-002-2	Sequence 2, Appl	934	5	10.9	7257	4	US-09-568-480-5	Sequence 5, Appl
862	5	10.9	1863	1	US-08-483-554B-2	Sequence 2, Appl	935	5	10.9	7257	4	US-09-568-472-5	Sequence 5, Appl
863	5	10.9	1863	1	US-08-798-691-2	Sequence 2, Appl	936	5	10.9	7257	4	US-09-568-472-5	Sequence 5, Appl
864	5	10.9	1863	1	US-08-798-691-4	Sequence 4, Appl	937	4	8.7	4	1	US-07-990-301A-10	Sequence 10, Appl
865	5	10.9	1863	1	US-08-798-691-6	Sequence 6, Appl	938	4	8.7	4	2	US-08-653-740-26	Sequence 26, Appl
866	5	10.9	1863	1	US-08-488-011B-2	Sequence 2, Appl	939	4	8.7	4	2	US-08-482-651-5	Sequence 5, Appl
867	5	10.9	1863	2	US-08-825-886-16	Sequence 16, Appl	940	4	8.7	4	2	US-08-803-899-25	Sequence 25, Appl
868	5	10.9	1863	2	US-08-603-753D-2	Sequence 2, Appl	941	4	8.7	4	2	US-09-073-594-26	Sequence 26, Appl
869	5	10.9	1863	3	US-08-825-487A-2	Sequence 2, Appl	942	4	8.7	4	3	US-09-275-925-26	Sequence 25, Appl
870	5	10.9	1863	3	US-08-825-487A-4	Sequence 4, Appl	943	4	8.7	4	4	US-08-435-568A-15	Sequence 15, Appl
871	5	10.9	1863	3	US-08-825-487A-6	Sequence 6, Appl	944	4	8.7	4	4	US-08-602-999A-104	Sequence 104, Appl
872	5	10.9	1863	4	US-09-074-476-2	Sequence 2, Appl	945	4	8.7	4	4	US-09-113-921-90	Sequence 90, Appl
873	5	10.9	1863	4	US-09-074-476-4	Sequence 4, Appl	946	4	8.7	4	4	US-09-113-921-91	Sequence 91, Appl
874	5	10.9	1863	4	US-09-074-476-6	Sequence 6, Appl	947	4	8.7	4	4	US-08-660-092-8	Sequence 8, Appl
875	5	10.9	1863	4	US-09-099-753-2	Sequence 2, Appl	948	4	8.7	4	4	US-09-198-723A-21	Sequence 21, Appl
876	5	10.9	1863	4	US-08-850-727-2	Sequence 2, Appl	949	4	8.7	4	4	US-09-066-481-20	Sequence 20, Appl
877	5	10.9	1863	4	US-08-986-106-2	Sequence 2, Appl	950	4	8.7	4	4	US-08-918-288-80	Sequence 80, Appl
878	5	10.9	1863	4	US-09-007-678B-49	Sequence 49, Appl	951	4	8.7	4	4	US-09-282-357-80	Sequence 80, Appl
879	5	10.9	1863	5	PCT-US95-10202-2	Sequence 2, Appl	952	4	8.7	4	4	US-08-278-865-104	Sequence 104, Appl
880	5	10.9	1863	5	PCT-US95-10203-2	Sequence 2, Appl	953	4	8.7	4	4	US-09-344-456-4	Sequence 4, Appl
881	5	10.9	1863	5	PCT-US95-10220-2	Sequence 2, Appl	954	4	8.7	5	1	US-07-956-848A-11	Sequence 1, Appl
882	5	10.9	1872	1	US-08-188-582-14	Sequence 14, Appl	955	4	8.7	5	1	US-07-956-848A-28	Sequence 28, Appl
883	5	10.9	1872	1	US-08-646-715-14	Sequence 14, Appl	956	4	8.7	5	1	US-07-956-848A-29	Sequence 29, Appl
884	5	10.9	1872	6	5386025-6	Patent No. 5386025	957	4	8.7	5	1	US-08-095-726-71	Sequence 71, Appl
885	5	10.9	1873	1	US-08-435-675B-4	Sequence 4, Appl	958	4	8.7	5	1	US-08-096-043-68	Sequence 68, Appl
886	5	10.9	1873	1	US-08-336-257A-7	Sequence 7, Appl	959	4	8.7	5	1	US-08-096-623A-76	Sequence 76, Appl
887	5	10.9	1893	1	US-08-188-582-11	Sequence 11, Appl	960	4	8.7	5	1	US-08-451-947-75	Sequence 75, Appl
888	5	10.9	1893	1	US-08-646-715-11	Sequence 11, Appl	961	4	8.7	5	1	US-08-477-727A-54	Sequence 54, Appl
889	5	10.9	1840	2	US-08-644-271-30	Sequence 30, Appl	962	4	8.7	5	1	US-08-471-956-1	Sequence 1, Appl
890	5	10.9	1964	2	US-08-790-912-3	Sequence 3, Appl	963	4	8.7	5	1	US-08-471-956-28	Sequence 28, Appl
891	5	10.9	2052	2	US-08-790-912-3	Sequence 2, Appl	964	4	8.7	5	1	US-08-471-956-29	Sequence 29, Appl
892	5	10.9	2231	1	US-08-153-799-16	Sequence 16, Appl	965	4	8.7	5	2	US-08-424-826A-75	Sequence 75, Appl
893	5	10.9	2285	4	US-09-308-375-2	Sequence 2, Appl	966	4	8.7	5	2	US-08-347-335A-8	Sequence 8, Appl
894	5	10.9	2289	3	US-09-051-019-2	Sequence 2, Appl	967	4	8.7	5	2	US-08-441-513A-8	Sequence 8, Appl
895	5	10.9	2324	1	US-08-283-857-1	Sequence 1, Appl	968	4	8.7	5	2	US-08-751-767A-25	Sequence 25, Appl
896	5	10.9	2324	6	PCT-US95-09819-1	Sequence 1, Appl	969	4	8.7	5	2	US-08-751-767A-27	Sequence 27, Appl
897	5	10.9	2327	5	5455158-1	Patent No. 5455158	970	4	8.7	5	3	US-08-928-694-75	Sequence 75, Appl
898	5	10.9	2386	2	US-09-016-366A-12	Sequence 12, Appl	971	4	8.7	5	3	US-08-855-925A-7	Sequence 7, Appl
899	5	10.9	2396	1	US-08-157-005-2	Sequence 2, Appl	972	4	8.7	5	3	US-08-581-662-10	Sequence 10, Appl
900	5	10.9	2396	4	US-08-747-863-2	Sequence 2, Appl	973	4	8.7	5	3	US-08-581-662-28	Sequence 28, Appl
901	5	10.9	2441	1	US-08-194-468-2	Sequence 2, Appl	974	4	8.7	5	4	US-09-194-613-11	Sequence 11, Appl
902	5	10.9	2441	3	US-08-961-739-2	Sequence 2, Appl	975	4	8.7	5	4	US-09-187-859-2113	Sequence 2113, Appl
903	5	10.9	2446	2	US-08-551-356-2	Sequence 2, Appl	976	4	8.7	5	4	US-09-187-859-2543	Sequence 2543, Appl

977 4 8.7 5 5 PCT-US91-06950-75 Sequence 75, Appl
978 4 8.7 6 1 US-08-095-726-61 Sequence 61, Appl
979 4 8.7 6 1 US-08-095-726-65 Sequence 65, Appl
980 4 8.7 6 1 US-08-096-043-58 Sequence 58, Appl
981 4 8.7 6 1 US-08-096-043-62 Sequence 62, Appl
982 4 8.7 6 1 US-08-269-257-12 Sequence 12, Appl
983 4 8.7 6 1 US-08-264-002-15 Sequence 15, Appl
984 4 8.7 6 1 US-07-890-422B-9 Sequence 9, Appl
985 4 8.7 6 1 US-07-890-422B-10 Sequence 10, Appl
986 4 8.7 6 1 US-07-890-422B-13 Sequence 13, Appl
987 4 8.7 6 1 US-07-890-422B-14 Sequence 14, Appl
988 4 8.7 6 1 US-08-096-623A-66 Sequence 66, Appl
989 4 8.7 6 1 US-08-096-623A-70 Sequence 70, Appl
990 4 8.7 6 1 US-08-451-947-76 Sequence 76, Appl
991 4 8.7 6 1 US-08-719-331-2 Sequence 2, Appl
992 4 8.7 6 2 US-08-724-548-50 Sequence 50, Appl
993 4 8.7 6 2 US-08-724-548-51 Sequence 51, Appl
994 4 8.7 6 2 US-08-424-826A-75 Sequence 75, Appl
995 4 8.7 6 2 US-08-463-667A-15 Sequence 15, Appl
996 4 8.7 6 2 US-08-465-380-18 Sequence 18, Appl
997 4 8.7 6 2 US-08-480-478-4 Sequence 4, Appl
998 4 8.7 6 2 US-08-486-397-18 Sequence 18, Appl
999 4 8.7 6 2 US-08-486-399-18 Sequence 18, Appl
1000 4 8.7 6 2 US-08-461-965-18 Sequence 18, Appl

ALIGNMENTS

RESULT 1
PCT-US93-03027-2
; Sequence 2, Application PC/TUS9303027
; GENERAL INFORMATION:
; APPLICANT: LEONARD, WARREN; TOLEDANO,
; APPLICANT: MICHEL
; TITLE OF INVENTION: CONTROL AND/OR
; TITLE OF INVENTION: PREVENTION OF BINDING OF NF- B/REL/DORSAL
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/03027
; FILING DATE: 19930401
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/862,987
; FILING DATE: 06-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUST
; REGISTRATION NUMBER: P-36,434
; REFERENCE/DOCKET NUMBER: 2026-4010 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-4800
; TELEFAX: 212-751-6849
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 558
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: mouse
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fibroblast
; CELL TYPE: NIH-3T3
; CELL LINE: NIH-3T3
; ORGANELLE:
; FEATURE:
; NAME/KEY: RelB
; LOCATION:
; IDENTIFICATION METHOD: contains region
; IDENTIFICATION METHOD: of high similarity with proteins of
; IDENTIFICATION METHOD: Rel family.
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: Ryseck, R.-P.; Bull, P.;
; AUTHORS: Takamiya, M.; Bours, V.; Siebenlist,
; AUTHORS: U.; Dobrzanski, P.; Bravo, R.
; TITLE: RelB, a new Rel family
; TITLE: transcription activator that can
; TITLE: interact with p50-NF- B.
; JOURNAL: Mol. Cell. Biol.
; VOLUME: 12
; ISSUE:
; PAGES: 674-684
; DATE: FEB-1992
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
PCT-US93-03027-2
Query Match 15.2%; Score 7; DB 5; Length 558;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 TDELEII 22
DB 33 TDELEII 39
RESULT 2
US-08-836-567-8
; Sequence 8, Application US/08836567
; Patent No. 6130367
; GENERAL INFORMATION:
; APPLICANT: Kossmann, Jens
; APPLICANT: Springer, Franziska
; APPLICANT: Abel, Gernot
; TITLE OF INVENTION: DNA MOLECULES THAT CODE FOR ENZYMES
; TITLE OF INVENTION: INVOLVED IN STARCH SYNTHESIS VECTORS BACTERIA TRANSGENIC
; TITLE OF INVENTION: PLANT CELLS AND PLANTS CONTAINING SAID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,567
; FILING DATE: 24-JUL-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: PCT/EP95/04415

Query Match 15.2%; Score 7; DB 4; Length 767;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 GSGSLS 30
DB 108 GSGSLS 114

RESULT 3
US-08-370-567-18
; Sequence 18, Application US/08370567
; Patent No. 5656726
; GENERAL INFORMATION:
; APPLICANT: Rosenberg, Steven
; APPLICANT: Doyle, Michael
; APPLICANT: Goodson, Robert
; TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/370,567
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER: US/08/061,514
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0941.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-370-567-18

Query Match 13.0%; Score 6; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 EIKTDE 18
DB 2 EIKTDE 7

RESULT 4
US-08-438-759-18
; Sequence 18, Application US/08438759
; Patent No. 5679782
; GENERAL INFORMATION:
; APPLICANT: Rosenberg, Steven
; APPLICANT: Doyle, Michael
; APPLICANT: Goodson, Robert
; TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/438,759
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: US/08/370,567
; FILING DATE:
; APPLICATION NUMBER: US/08/061,514
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0941.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; US-08-438-759-18

Query Match 13.0%; Score 6; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 EIKTDE 18
DB 2 EIKTDE 7

RESULT 5
PCT-US94-05684-18
; Sequence 18, Application PC/TUS9405684
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor

Query Match 13.0%; Score 6; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 EIKTDE 18
DB 2 EIKTDE 7

RESULT 5
PCT-US94-05684-18
; Sequence 18, Application PC/TUS9405684
; GENERAL INFORMATION:
; APPLICANT: Chiron Corporation
; TITLE OF INVENTION: Peptide Inhibitors of Urokinase Receptor

```

; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05684
; FILING DATE: 19 MAY 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0941.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEetical: NO
PCT-US94-05684-18

Query Match 13.0%; Score 6; DB 5; Length 17;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 EIKTDE 18
DB 2 EIKTDE 7

```

```

; TITLE OF INVENTION: Activity
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05684
; FILING DATE: 19 MAY 1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Grant D.
; REGISTRATION NUMBER: 31,259
; REFERENCE/DOCKET NUMBER: 0941.100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2706
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEtical: NO
PCT-US94-05684-18

Query Match 13.0%; Score 6; DB 2; Length 114;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 IGGSGS 28
DB 51 IGGSGS 56

RESULT 7
US-09-367-953B-120
; Sequence 120, Application US/09367953B
; Patent No. 6287572
; GENERAL INFORMATION:
; APPLICANT: Kingsman, Alan J
; APPLICANT: Kingsman, Susan M
; TITLE OF INVENTION: Anti-HIV peptides and proteins
; FILE REFERENCE: 550-154
; CURRENT APPLICATION NUMBER: US/09/367,953B
; CURRENT FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: PCT/GB98/00563
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: GB 9703802.0
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 134
; TYPE: PPT
; ORGANISM: Human immunodeficiency virus type 2
US-09-367-953B-120

Query Match 13.0%; Score 6; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 KTDELE 20
DB 12 KTDELE 17

RESULT 8
US-09-367-953B-121
; Sequence 121, Application US/09367953B
; Patent No. 6287572
; GENERAL INFORMATION:
; APPLICANT: Kingsman, Alan J
; APPLICANT: Kingsman, Susan M
; TITLE OF INVENTION: Anti-HIV peptides and proteins
; FILE REFERENCE: 550-154
; CURRENT APPLICATION NUMBER: US/09/367,953B
; CURRENT FILING DATE: 1999-08-24
; PRIOR APPLICATION NUMBER: PCT/GB98/00563
; PRIOR FILING DATE: 1998-02-23
; PRIOR APPLICATION NUMBER: GB 9703802.0
; PRIOR FILING DATE: 1997-02-24
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 121
; LENGTH: 134

```

```
; TYPE: PRT
; ORGANISM: Human immunodeficiency virus type 2
US-09-367-953B-121

Query Match      13.0%; Score 6; DB 4; Length 134;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 KTDELE 20
Db 12 KTDELE 17

RESULT 9
US-08-482-918-56
; Sequence 56, Application US/08482918
; Patent No. 6207417
; GENERAL INFORMATION:
; APPLICANT: Zsebo, Kristina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Martin, Francis H.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,681
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 01017/33005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-918-56

Query Match      13.0%; Score 6; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 FFRLEN 37
Db 139 FFRLEN 144

RESULT 10
US-09-224-681-56
; Sequence 56, Application US/09224681
; Patent No. 6207454
; GENERAL INFORMATION:
; APPLICANT: Zsebo, Kristina M.
; APPLICANT: Bosselman, Robert A.
```

Patent No. 6207802
GENERAL INFORMATION:
APPLICANT: Zsebo, Krisztina M.
APPLICANT: Bosseman, Robert A.
APPLICANT: Sugds, Sidney V.
APPLICANT: Martin, Francis H.
TITLE OF INVENTION: Stem Cell Factor
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,728A
FILING DATE: 09-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/982,255
FILING DATE: 25-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/589,701
FILING DATE: 01-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/573,616
FILING DATE: 24-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/537,198
FILING DATE: 11-JUN-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/422,383
FILING DATE: 16-OCT-1989
ATTORNEY/AGENT INFORMATION:
NAME: Clough, David W.
REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/32956
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-728A-56

Query Match 13.0%; Score 6; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 FFRFLN 37
Db 139 FFRFLN 144

RESULT 12
US-09-383-586-33
Sequence 33, Application US/09383586
Patent No. 6242419
GENERAL INFORMATION:
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Abernethy, Nevin

APPLICANT: Onrust, Rene
APPLICANT: Kumble, Anand
APPLICANT: Murison, Greg
TITLE OF INVENTION: Compounds isolated from stromal cells
TITLE OF INVENTION: and methods for their use
FILE REFERENCE: 11000.1037c1
CURRENT APPLICATION NUMBER: US/09/383.586
CURRENT FILING DATE: 1999-08-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 322
TYPE: PRT
ORGANISM: Human
US-09-383-586-33

Query Match 13.0%; Score 6; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TSLKLN 9
Db 203 TSLKLN 208

RESULT 13
US-09-079-415-2
Sequence 2, Application US/09079415
Patent No. 6013452
GENERAL INFORMATION:
APPLICANT: Christensen, Tove
APPLICANT: Lehmebeck, Jan
TITLE OF INVENTION: A Fungus Wherein The area, pepC and/or
TITLE OF INVENTION: pepe Genes Have Been Inactivated
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6013452o No. 6013452disk of No. 6013452th America, Inc.
STREET: 405 Lexington Avenue
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079.415
FILING DATE: 14-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Rozek, Carol
REGISTRATION NUMBER: 36,993
REFERENCE/DOCKET NUMBER: 4657.204-US
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-079-415-2

Query Match 13.0%; Score 6; DB 3; Length 397;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSLN 30
Db 146 GSGSLN 151

RESULT 14

US-08-328-314-2
; Sequence 2, Application US/08328314
; Patent No. 5674728
; GENERAL INFORMATION:
; APPLICANT: Buxton, Frank
; APPLICANT: Jaral, Gabor
; APPLICANT: Visser, Jacob
; TITLE OF INVENTION: Fungal Protease
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII editor
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,314
; FILING DATE: TBA
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: 4-19746/A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 541-8615
; TELEFAX: (919) 541-8689
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-328-314-2

Query Match 13.0%; Score 6; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSLs 30
Db 147 GSGSLs 152

RESULT 15

US-08-731-045-2
; Sequence 2, Application US/08731045
; Patent No. 5756338
; GENERAL INFORMATION:
; APPLICANT: Buxton, Frank
; APPLICANT: Jaral, Gabor
; APPLICANT: Visser, Jacob
; TITLE OF INVENTION: Fungal Protease
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY CORPORATION
; STREET: 520 White Plains Road
; CITY: Tarrytown
; STATE: New York
; COUNTRY: USA
; ZIP: 10591-9725
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII Editor
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/731,045
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/328,314
FILING DATE: October 24, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Ferrara, Gregory D.
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 4-19746/A/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-3318
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-731-045-2

Query Match 13.0%; Score 6; DB 1; Length 398;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 GSGSLs 30
Db 147 GSGSLs 152

Search completed: November 5, 2002, 11:06:56
Job time : 38.1045 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 11:05:33 ; Search time 144.179 seconds
(without alignments)
112.299 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 46

Sequence: I MKRTLSLKNDKFEIKTDELE.....GSLSTFFRLFNKSFQALGK 46

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	46	22	US-09-833-017-2
2	46	100.0	46	22	US-09-833-017-8
3	46	100.0	46	22	US-09-833-017-9
4	46	100.0	46	22	US-09-833-017-11
5	46	100.0	46	22	US-09-833-017-12
6	46	100.0	46	22	US-09-833-017-13
7	41	89.1	46	22	US-09-833-017-7
					Sequence 2, Appli
					Sequence 8, Appli
					Sequence 9, Appli
					Sequence 11, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 7, Appli

58.7	27	8	43	22	US-09-833-017-10	Sequence 10, Appli
45.7	21	9	21	22	US-09-833-017-14	Sequence 14, Appli
15.2	7	10	46	24	US-10-029-386-31630	Sequence 31630, A
15.2	7	11	118	23	US-09-902-540-12650	Sequence 12650, A
15.2	7	12	183	21	US-09-708-427-66071	Sequence 66071, A
15.2	7	13	229	21	US-09-708-427-66070	Sequence 66070, A
15.2	7	14	235	21	US-09-708-427-66069	Sequence 66069, A
15.2	7	15	352	1	PCT-US01-14827-12064	Sequence 12064, A
15.2	7	16	353	21	US-09-708-427-83339	Sequence 83339, A
15.2	7	17	393	26	US-60-167-217-1542	Sequence 1542, Ap
15.2	7	18	393	26	US-60-173-464-1254	Sequence 1254, Ap
15.2	7	19	397	23	US-09-902-540-16267	Sequence 16267, A
15.2	7	20	427	21	US-09-708-427-83338	Sequence 83338, A
15.2	7	21	430	21	US-09-708-427-83337	Sequence 83337, A
15.2	7	22	447	17	US-09-328-352-7917	Sequence 7917, Ap
15.2	7	23	457	16	US-09-252-991A-30850	Sequence 30850, A
15.2	7	24	558	3	US-07-862-987A-2	Sequence 2, Appli
15.2	7	25	586	21	US-09-708-427-5170	Sequence 5170, Ap
15.2	7	26	612	21	US-09-708-427-5169	Sequence 5169, Ap
15.2	7	27	618	21	US-09-708-427-5168	Sequence 5168, Ap
15.2	7	28	658	1	PCT-US01-08631-48340	Sequence 48340, A
15.2	7	29	767	20	US-09-606-304-8	Sequence 8, Appli
15.2	7	30	767	20	US-09-614-150-1512	Sequence 1512, Ap
15.2	7	31	767	26	US-60-191-637-1515	Sequence 1515, Ap
15.2	7	32	767	26	US-60-191-681-1203	Sequence 1203, Ap
15.2	7	33	929	20	US-09-614-150-39363	Sequence 39363, A
15.2	7	34	929	26	US-60-173-464-29732	Sequence 29732, A
15.2	7	35	929	26	US-60-191-637-38989	Sequence 38989, A
15.2	7	36	929	26	US-60-191-681-30211	Sequence 30211, A
15.2	7	37	1357	16	US-09-270-767-44971	Sequence 44971, A
15.2	7	38	1394	16	US-09-248-796-19555	Sequence 19555, A
15.2	7	39	1394	26	US-60-096-409-19555	Sequence 19555, A
15.2	7	40	1767	26	US-60-230-445-1919	Sequence 1919, Ap
15.2	7	41	1773	26	US-60-230-445-1433	Sequence 1433, Ap
15.2	7	42	2771	20	US-09-614-150-18432	Sequence 18432, A
15.2	7	43	2771	26	US-60-167-217-18522	Sequence 18522, A
15.2	7	44	2771	26	US-60-173-464-15176	Sequence 15176, A
15.2	7	45	2771	26	US-60-191-637-18488	Sequence 18488, A
15.2	7	46	2771	26	US-60-191-681-14627	Sequence 14627, A
15.2	7	47	576775	12	US-08-895-611-2	Sequence 2, Appli
15.2	7	48	576775	12	US-08-895-611D-2	Sequence 2, Appli
15.2	7	49	576775	12	US-09-895-611D-2	Sequence 21, Appli
13.0	6	50	19	19	US-09-549-676A-21	Sequence 472, App
13.0	6	51	20	22	US-09-895-828-472	Sequence 473, App
13.0	6	52	20	22	US-09-895-828-473	Sequence 118541, A
13.0	6	53	23	21	US-09-724-059-118541	Sequence 118541, A
13.0	6	54	23	21	US-09-724-059-118730	Sequence 118730, A
13.0	6	55	23	21	US-09-724-059-118910	Sequence 118910, A
13.0	6	56	23	21	US-09-724-059-119099	Sequence 119099, A
13.0	6	57	23	21	US-09-724-059-119288	Sequence 119288, A
13.0	6	58	23	21	US-09-724-059-119459	Sequence 119459, A
13.0	6	59	23	21	US-09-724-059-119666	Sequence 119666, A
13.0	6	60	23	21	US-09-724-059-119837	Sequence 119837, A
13.0	6	61	23	21	US-09-724-059-120017	Sequence 120017, A
13.0	6	62	23	21	US-09-724-059-120206	Sequence 120206, A
13.0	6	63	23	21	US-09-724-059-120428	Sequence 120428, A
13.0	6	64	23	21	US-09-724-059-120806	Sequence 120806, A
13.0	6	65	23	21	US-09-724-059-121166	Sequence 121166, A
13.0	6	66	23	21	US-09-724-059-121544	Sequence 121544, A
13.0	6	67	23	21	US-09-724-059-121922	Sequence 121922, A
13.0	6	68	23	21	US-09-724-059-122264	Sequence 122264, A
13.0	6	69	23	21	US-09-724-059-122678	Sequence 122678, A
13.0	6	70	23	21	US-09-724-059-123020	Sequence 123020, A
13.0	6	71	23	21	US-09-724-059-123380	Sequence 123380, A
13.0	6	72	23	21	US-09-724-059-123758	Sequence 123758, A
13.0	6	73	23	21	US-09-724-059-127134	Sequence 127134, A
13.0	6	74	23	21	US-09-724-059-127260	Sequence 127260, A
13.0	6	75	23	21	US-09-724-059-127380	Sequence 127380, A
13.0	6	76	23	21	US-09-724-059-127506	Sequence 127506, A
13.0	6	77	23	21	US-09-724-059-127632	Sequence 127632, A
13.0	6	78	23	21	US-09-724-059-127746	Sequence 127746, A
13.0	6	79	23	21	US-09-724-059-127884	Sequence 127884, A
13.0	6	80	23	21	US-09-724-059-127998	Sequence 127998, A

81	6	13.0	23	21	US-09-724-059-128118	Sequence 128118, A	154	6	13.0	67	26	US-60-177-571-4860	Sequence 4860, Ap
82	6	13.0	23	21	US-09-724-059-128244	Sequence 128244, A	155	6	13.0	69	1	PCT-US01-00863-31338	Sequence 31338, A
83	6	13.0	23	21	US-09-724-059-128392	Sequence 128392, A	156	6	13.0	69	1	PCT-US01-14827-13422	Sequence 13422, A
84	6	13.0	23	21	US-09-724-059-128644	Sequence 128644, A	157	6	13.0	69	22	US-09-864-761-38134	Sequence 38134, A
85	6	13.0	23	21	US-09-724-059-128884	Sequence 128884, A	158	6	13.0	69	26	US-60-236-359-20765	Sequence 20765, A
86	6	13.0	23	21	US-09-724-059-129136	Sequence 129136, A	159	6	13.0	70	1	PCT-US01-01324-2296	Sequence 2296, Ap
87	6	13.0	23	21	US-09-724-059-129388	Sequence 129388, A	160	6	13.0	70	1	PCT-US01-01351-230	Sequence 230, App
88	6	13.0	23	21	US-09-724-059-129616	Sequence 129616, A	161	6	13.0	70	21	US-09-764-887-230	Sequence 230, App
89	6	13.0	23	21	US-09-724-059-129892	Sequence 129892, A	162	6	13.0	70	24	US-10-073-961-230	Sequence 230, App
90	6	13.0	23	21	US-09-724-059-130120	Sequence 130120, A	163	6	13.0	70	24	US-10-079-979-2296	Sequence 2296, Ap
91	6	13.0	23	21	US-09-724-059-130360	Sequence 130360, A	164	6	13.0	71	26	US-60-186-661-437	Sequence 606, App
92	6	13.0	23	21	US-09-724-059-130612	Sequence 130612, A	165	6	13.0	71	26	US-60-186-661-437	Sequence 606, App
93	6	13.0	23	21	US-09-724-059-130327	Sequence 130327, A	166	6	13.0	72	15	US-09-134-000-4826	Sequence 4826, Ap
94	6	13.0	23	21	US-09-724-059-140587	Sequence 140587, A	167	6	13.0	72	15	US-09-134-000-4826	Sequence 4826, Ap
95	6	13.0	23	21	US-09-724-059-141787	Sequence 141787, A	168	6	13.0	74	21	US-09-739-449-10367	Sequence 10367, A
96	6	13.0	23	21	US-09-724-059-143047	Sequence 143047, A	169	6	13.0	74	22	US-09-803-110-10267	Sequence 10267, A
97	6	13.0	23	21	US-09-724-059-144307	Sequence 144307, A	170	6	13.0	75	26	US-60-169-840-4921	Sequence 4921, Ap
98	6	13.0	23	21	US-09-724-059-145447	Sequence 145447, A	171	6	13.0	76	18	US-09-489-039A-7560	Sequence 7560, Ap
99	6	13.0	23	21	US-09-724-059-146827	Sequence 146827, A	172	6	13.0	78	1	PCT-US01-01354-16970	Sequence 16970, A
100	6	13.0	23	21	US-09-724-059-146827	Sequence 146827, A	173	6	13.0	78	21	US-09-764-905-16970	Sequence 16970, A
101	6	13.0	23	21	US-09-724-059-149167	Sequence 149167, A	174	6	13.0	79	1	PCT-US01-05109-964	Sequence 964, App
102	6	13.0	23	21	US-09-724-059-150427	Sequence 150427, A	175	6	13.0	82	21	US-09-760-476-1727	Sequence 1727, Ap
103	6	13.0	23	21	US-09-724-059-151907	Sequence 151907, A	176	6	13.0	83	1	PCT-US01-01334-5236	Sequence 5236, Ap
104	6	13.0	23	21	US-09-724-059-154427	Sequence 154427, A	177	6	13.0	83	21	US-09-764-874-5236	Sequence 5236, Ap
105	6	13.0	23	21	US-09-724-059-156827	Sequence 156827, A	178	6	13.0	85	16	US-09-270-767-41443	Sequence 41443, A
106	6	13.0	23	21	US-09-724-059-159347	Sequence 159347, A	179	6	13.0	86	22	US-09-836-734-190	Sequence 190, App
107	6	13.0	23	21	US-09-724-059-161867	Sequence 161867, A	180	6	13.0	86	26	US-60-233-940-132	Sequence 132, App
108	6	13.0	23	21	US-09-724-059-164147	Sequence 164147, A	181	6	13.0	88	26	US-60-186-661-434	Sequence 434, App
109	6	13.0	23	21	US-09-724-059-166907	Sequence 166907, A	182	6	13.0	88	26	US-60-192-737-1133	Sequence 1133, Ap
110	6	13.0	23	21	US-09-724-059-169187	Sequence 169187, A	183	6	13.0	88	26	US-60-194-106-698	Sequence 698, App
111	6	13.0	23	21	US-09-724-059-171587	Sequence 171587, A	184	6	13.0	89	26	US-60-186-661-440	Sequence 440, App
112	6	13.0	23	21	US-09-724-059-174107	Sequence 174107, A	185	6	13.0	93	16	US-09-270-767-57065	Sequence 57065, A
113	6	13.0	23	26	PCT-US99-03939-164	Sequence 840, App	186	6	13.0	93	16	US-09-270-767-57065	Sequence 57065, A
114	6	13.0	27	1	PCT-US99-03939-164	Sequence 164, App	187	6	13.0	96	26	US-60-192-737-1113	Sequence 1113, Ap
115	6	13.0	30	1	PCT-US01-00663-27981	Sequence 27981, A	188	6	13.0	96	26	US-60-194-106-688	Sequence 688, App
116	6	13.0	30	22	US-09-864-761-34974	Sequence 34974, A	189	6	13.0	100	26	US-60-178-305-929	Sequence 929, App
117	6	13.0	30	26	US-60-236-359-16494	Sequence 16494, A	190	6	13.0	100	26	US-60-178-305-931	Sequence 931, App
118	6	13.0	34	1	PCT-US01-01334-18951	Sequence 18951, A	191	6	13.0	101	1	PCT-US00-05988-1709	Sequence 1709, Ap
119	6	13.0	34	21	US-09-764-905-18951	Sequence 18951, A	192	6	13.0	101	23	US-09-925-300-1709	Sequence 1709, Ap
120	6	13.0	34	21	US-10-092-399-18951	Sequence 18951, A	193	6	13.0	103	26	US-60-139-669-485	Sequence 485, App
121	6	13.0	34	26	US-60-163-233-2268	Sequence 2268, Ap	194	6	13.0	105	1	PCT-US01-08631-44150	Sequence 44150, A
122	6	13.0	35	3	US-07-906-308-10	Sequence 10, Appl	195	6	13.0	105	17	US-09-389-956-70	Sequence 70, Appl
123	6	13.0	35	3	US-07-906-308-10	Sequence 33, Appl	196	6	13.0	105	20	US-09-675-784A-12517	Sequence 12517, A
124	6	13.0	40	26	US-60-162-243-477	Sequence 477, App	197	6	13.0	110	13	US-08-905-135-467	Sequence 467, App
125	6	13.0	40	26	US-60-169-840-8280	Sequence 8280, Ap	198	6	13.0	112	19	US-09-583-110-3690	Sequence 3690, Ap
126	6	13.0	42	26	US-60-195-300-15	Sequence 15, Appl	199	6	13.0	114	15	US-09-150-816-52	Sequence 52, Appl
127	6	13.0	43	26	US-60-195-300-15	Sequence 15, Appl	200	6	13.0	115	26	US-60-178-305-1127	Sequence 1127, Ap
128	6	13.0	43	26	US-60-182-902-608	Sequence 608, App	201	6	13.0	116	15	US-09-107-433-4238	Sequence 4238, Ap
129	6	13.0	45	26	US-60-198-954-608	Sequence 608, App	202	6	13.0	116	26	US-60-182-467-1968	Sequence 1968, Ap
130	6	13.0	48	11	US-08-756-693-5	Sequence 5, Appl	203	6	13.0	119	16	US-09-270-767-44716	Sequence 44716, A
131	6	13.0	50	11	US-08-756-693A-5	Sequence 5, Appl	204	6	13.0	121	13	US-08-993-002A-4996	Sequence 4996, Ap
132	6	13.0	51	7	US-08-330-433-9	Sequence 9, Appl	205	6	13.0	121	17	US-09-330-337-321	Sequence 321, App
133	6	13.0	51	7	US-08-330-433-9	Sequence 11, Appl	206	6	13.0	121	18	US-09-417-507-41433	Sequence 41433, A
134	6	13.0	53	1	PCT-US99-3283B-2932	Sequence 2932, Ap	207	6	13.0	121	26	US-60-182-567-612	Sequence 612, App
135	6	13.0	53	1	PCT-US00-46524B-7767	Sequence 7767, Ap	208	6	13.0	122	1	PCT-US99-03939-158	Sequence 158, App
136	6	13.0	59	17	US-09-345-464-39	Sequence 39, Appl	209	6	13.0	122	13	US-08-993-002A-4997	Sequence 4997, Ap
137	6	13.0	59	17	US-09-345-464-39	Sequence 39, Appl	210	6	13.0	122	16	US-09-248-796-24846	Sequence 24846, A
138	6	13.0	59	20	US-09-606-317-39	Sequence 39, Appl	211	6	13.0	122	21	US-09-716-129-162	Sequence 162, App
139	6	13.0	59	22	US-09-834-366-23839	Sequence 23839, A	212	6	13.0	122	23	US-09-902-540-11364	Sequence 11364, A
140	6	13.0	59	22	US-09-834-366-23839	Sequence 23839, A	213	6	13.0	122	26	US-60-096-409-24846	Sequence 24846, A
141	6	13.0	60	26	US-60-197-873-23839	Sequence 23839, A	214	6	13.0	123	26	US-60-178-305-1131	Sequence 1131, A
142	6	13.0	62	7	US-60-196-712-2630	Sequence 2630, Ap	215	6	13.0	124	26	US-60-192-737-1114	Sequence 1114, Ap
143	6	13.0	62	7	US-60-330-433-10	Sequence 10, Appl	216	6	13.0	124	26	US-60-194-106-689	Sequence 689, App
144	6	13.0	62	21	US-09-758-472-5612	Sequence 5612, Ap	217	6	13.0	124	26	US-60-178-305-930	Sequence 930, App
145	6	13.0	62	26	US-60-178-310-333	Sequence 333, App	218	6	13.0	129	26	US-08-993-002A-8527	Sequence 8527, Ap
146	6	13.0	63	16	US-09-252-691C-10974	Sequence 481, App	219	6	13.0	130	13	US-08-993-002A-8526	Sequence 8526, Ap
147	6	13.0	63	16	US-09-252-691C-10974	Sequence 10974, A	220	6	13.0	131	13	PCT-US01-01347-346	Sequence 346, App
148	6	13.0	66	1	PCT-US01-08631-35062	Sequence 10974, A	221	6	13.0	132	8	US-08-487-032A-790	Sequence 790, App
149	6	13.0	66	1	PCT-US99-2283B-2931	Sequence 35062, A	222	6	13.0	132	9	US-08-561-469A-790	Sequence 790, App
150	6	13.0	68	7	US-08-330-433-12	Sequence 2931, Ap	223	6	13.0	132	11	US-08-761-184-1029	Sequence 1029, Ap
151	6	13.0	68	7	US-08-330-433-12	Sequence 12, Appl	224	6	13.0	132	12	US-08-821-931-1029	Sequence 1029, Ap
152	6	13.0	68	7	US-08-330-433-12	Sequence 264, App	225	6	13.0	132	12	US-08-821-931-1029	Sequence 1029, Ap
153	6	13.0	68	7	US-08-330-433-12	Sequence 4814, App	226	6	13.0	133	1	PCT-US01-14827-15643	Sequence 15643, A

227	6	13.0	134	1	PCT-US01-14827-15658	Sequence 15658, A	300	6	13.0	186	26	US-60-150-584-1011	Sequence 1011, Ap
228	6	13.0	135	26	US-60-182-567-614	Sequence 614, Ap	301	6	13.0	186	26	US-60-192-737-1118	Sequence 1118, Ap
229	6	13.0	136	26	US-60-188-162-3521	Sequence 3521, Ap	302	6	13.0	186	26	US-60-194-106-693	Sequence 693, App
230	6	13.0	137	21	US-09-733-089-14808	Sequence 14808, A	303	6	13.0	191	26	US-60-195-136-825	Sequence 825, App
231	6	13.0	138	22	US-09-816-660-14808	Sequence 14808, A	304	6	13.0	195	26	US-60-213-845-344	Sequence 344, App
232	6	13.0	139	17	US-09-358-322-11	Sequence 11, Appl	305	6	13.0	197	16	US-09-252-991A-28165	Sequence 28165, Ap
233	6	13.0	139	17	US-09-358-322A-11	Sequence 11, Appl	306	6	13.0	197	16	US-60-178-305-1128	Sequence 1128, Ap
234	6	13.0	139	24	US-10-004-494A-11	Sequence 11, Appl	307	6	13.0	198	16	US-09-270-767-61873	Sequence 61873, A
235	6	13.0	139	26	US-60-207-213-2711	Sequence 2711, App	308	6	13.0	198	16	US-09-270-849B-193908	Sequence 193908, A
236	6	13.0	141	15	US-09-198-452A-1236	Sequence 1236, Ap	309	6	13.0	198	21	US-09-708-427-57479	Sequence 57479, A
237	6	13.0	147	1	PCT-US01-05363-7	Sequence 7, Appli	310	6	13.0	198	21	US-09-708-427-57479	Sequence 57479, A
238	6	13.0	147	21	US-09-708-427-14541	Sequence 14541, A	311	6	13.0	201	26	US-60-316-362-3153	Sequence 3153, Ap
239	6	13.0	147	21	US-09-725-413A-4	Sequence 4, Appli	312	6	13.0	203	19	US-09-583-110-3736	Sequence 3736, Ap
240	6	13.0	147	21	US-09-789-836-7	Sequence 7, Appli	313	6	13.0	206	26	US-60-167-217-12329	Sequence 12329, A
241	6	13.0	147	21	US-09-134-001C-4679	Sequence 4679, Ap	314	6	13.0	208	21	US-09-760-475-1934	Sequence 1934, Ap
242	6	13.0	148	15	US-09-338-185A-956	Sequence 956, App	315	6	13.0	208	21	US-09-760-475-1934	Sequence 1934, Ap
243	6	13.0	148	16	US-09-270-767-38744	Sequence 38744, A	316	6	13.0	209	22	US-09-870-759-39	Sequence 39, Appl
244	6	13.0	148	16	US-09-270-767-53961	Sequence 53961, A	317	6	13.0	209	22	US-09-870-759-39	Sequence 39, Appl
245	6	13.0	148	16	US-09-270-849B-189782	Sequence 189782, A	318	6	13.0	212	1	PCT-US01-08631-55189	Sequence 55189, A
246	6	13.0	148	18	US-09-450-969-6155	Sequence 6155, Ap	319	6	13.0	214	15	US-09-107-433-3882	Sequence 3882, Ap
247	6	13.0	148	18	US-09-450-969-7130	Sequence 7130, Ap	320	6	13.0	214	26	US-60-152-737-1120	Sequence 1120, Ap
248	6	13.0	149	1	PCT-US01-08631-47633	Sequence 47633, A	321	6	13.0	214	26	US-60-194-106-695	Sequence 695, App
249	6	13.0	150	1	PCT-US01-01341-1052	Sequence 1052, Ap	322	6	13.0	216	26	US-60-192-737-1119	Sequence 1119, Ap
250	6	13.0	150	18	US-09-438-185-956	Sequence 956, App	323	6	13.0	216	26	US-60-194-106-694	Sequence 694, App
251	6	13.0	150	18	US-09-438-185A-956	Sequence 956, App	324	6	13.0	219	26	US-60-178-305-1130	Sequence 1130, Ap
252	6	13.0	150	24	US-09-764-864-1052	Sequence 1052, Ap	325	6	13.0	220	1	PCT-US98-06371-546	Sequence 546, App
253	6	13.0	150	24	US-10-080-129-1052	Sequence 1052, Ap	326	6	13.0	220	12	US-08-881-227-176	Sequence 176, App
254	6	13.0	150	26	US-60-141-862-678	Sequence 678, App	327	6	13.0	220	12	US-08-881-227-176	Sequence 176, App
255	6	13.0	150	26	US-60-141-863-678	Sequence 678, App	328	6	13.0	220	12	US-08-881-227-176	Sequence 176, App
256	6	13.0	151	16	US-09-270-767-42527	Sequence 42527, A	329	6	13.0	221	26	US-08-993-002A-5417	Sequence 5417, Ap
257	6	13.0	151	16	US-09-270-767-42527	Sequence 42527, A	330	6	13.0	221	26	US-60-146-394-1698	Sequence 1698, Ap
258	6	13.0	152	26	US-60-170-374-3444	Sequence 3444, Ap	331	6	13.0	221	26	US-60-147-189-977	Sequence 977, App
259	6	13.0	158	26	US-60-178-305-1125	Sequence 1125, Ap	332	6	13.0	222	26	US-60-178-305-1129	Sequence 1129, Ap
260	6	13.0	160	16	US-09-270-767-36332	Sequence 36332, A	333	6	13.0	222	26	US-09-201-228A-1005	Sequence 1005, Ap
261	6	13.0	160	16	US-09-270-767-51549	Sequence 51549, A	334	6	13.0	223	21	US-09-708-427-21929	Sequence 21929, A
262	6	13.0	160	16	US-09-270-849B-184782	Sequence 184782, A	335	6	13.0	223	21	US-09-708-427-21929	Sequence 21929, A
263	6	13.0	160	21	US-09-489-039A-8791	Sequence 8791, Ap	336	6	13.0	224	26	US-09-614-150-12297	Sequence 12297, A
264	6	13.0	160	21	US-09-614-450-701	Sequence 701, App	337	6	13.0	224	26	US-60-185-360-244	Sequence 244, App
265	6	13.0	161	18	US-09-708-427-60528	Sequence 60528, A	338	6	13.0	224	26	US-60-186-661-435	Sequence 435, App
266	6	13.0	162	26	US-09-417-507-37306	Sequence 37306, A	339	6	13.0	224	26	US-60-186-661-438	Sequence 438, App
267	6	13.0	162	26	US-60-194-106-691	Sequence 691, App	340	6	13.0	225	11	US-60-186-661-438	Sequence 438, App
268	6	13.0	163	11	US-08-792-511-17	Sequence 17, Appl	341	6	13.0	225	11	US-08-756-693-8	Sequence 8, Appli
269	6	13.0	163	16	US-09-252-691-5947	Sequence 5947, Ap	342	6	13.0	225	26	US-08-756-693A-8	Sequence 8, Appli
270	6	13.0	163	16	US-09-252-691C-5947	Sequence 5947, Ap	343	6	13.0	226	15	US-60-186-661-436	Sequence 436, App
271	6	13.0	166	16	US-09-270-767-33483	Sequence 33483, A	344	6	13.0	226	15	US-60-243-468-1247	Sequence 1247, Ap
272	6	13.0	166	16	US-09-270-767-48700	Sequence 48700, A	345	6	13.0	226	15	US-09-107-532-3948	Sequence 3948, Ap
273	6	13.0	166	16	US-09-270-767-48700	Sequence 48700, A	346	6	13.0	226	15	US-09-107-532A-3948	Sequence 3948, Ap
274	6	13.0	167	26	US-60-192-737-1117	Sequence 1117, Ap	347	6	13.0	228	26	US-60-324-109-27251	Sequence 27251, A
275	6	13.0	167	26	US-60-194-106-692	Sequence 692, App	348	6	13.0	231	21	US-09-733-089-1021	Sequence 1021, Ap
276	6	13.0	168	26	US-60-192-737-1121	Sequence 1121, Ap	349	6	13.0	231	22	US-09-816-660-1021	Sequence 1021, Ap
277	6	13.0	168	26	US-60-194-106-696	Sequence 696, App	350	6	13.0	233	26	US-09-708-427-21928	Sequence 21928, A
278	6	13.0	169	26	US-60-182-567-613	Sequence 613, App	351	6	13.0	233	26	US-60-312-544-6615	Sequence 6615, Ap
279	6	13.0	171	26	US-60-192-737-1122	Sequence 1122, Ap	352	6	13.0	236	26	US-09-902-540-15029	Sequence 15029, A
280	6	13.0	171	26	US-60-194-106-697	Sequence 697, App	353	6	13.0	238	26	US-60-185-360-246	Sequence 246, App
281	6	13.0	172	21	PCT-US01-01341-1315	Sequence 1315, Ap	354	6	13.0	242	26	US-60-128-476-2875	Sequence 2875, Ap
282	6	13.0	172	21	US-09-764-864-1315	Sequence 1315, Ap	355	6	13.0	243	15	US-09-198-452A-284	Sequence 284, App
283	6	13.0	172	24	US-10-080-129-1315	Sequence 1315, Ap	356	6	13.0	243	21	US-09-708-427-57477	Sequence 57477, A
284	6	13.0	174	13	US-08-625-811-1685	Sequence 1685, Ap	357	6	13.0	243	21	US-09-760-475-2085	Sequence 2085, Ap
285	6	13.0	174	13	US-08-993-002A-5416	Sequence 5416, Ap	358	6	13.0	245	18	US-09-438-185-273	Sequence 273, App
286	6	13.0	174	20	US-09-614-450-700	Sequence 700, App	359	6	13.0	247	19	US-09-543-681A-5522	Sequence 5522, Ap
287	6	13.0	174	21	US-09-708-427-60527	Sequence 60527, A	360	6	13.0	248	26	US-60-167-217-7825	Sequence 7825, Ap
288	6	13.0	175	26	US-60-186-661-439	Sequence 439, App	361	6	13.0	248	26	US-60-173-464-6250	Sequence 6250, Ap
289	6	13.0	178	26	US-60-178-305-1126	Sequence 1126, Ap	362	6	13.0	248	26	US-60-285-697-34	Sequence 34, Appl
290	6	13.0	180	11	US-08-756-693-2	Sequence 2, Appli	363	6	13.0	248	26	US-60-285-697-34	Sequence 34, Appl
291	6	13.0	180	11	US-08-756-693A-2	Sequence 2, Appli	364	6	13.0	252	1	PCT-US01-08631-44716	Sequence 44716, A
292	6	13.0	183	26	US-60-192-737-1115	Sequence 1115, Ap	365	6	13.0	252	21	PCT-US01-18569-2358	Sequence 2358, Ap
293	6	13.0	183	26	US-60-194-106-690	Sequence 690, App	366	6	13.0	252	21	US-09-760-475-2092	Sequence 2092, Ap
294	6	13.0	184	26	US-60-142-896-1235	Sequence 1235, Ap	367	6	13.0	252	1	PCT-US01-08631-43146	Sequence 43146, A
295	6	13.0	185	26	US-60-161-932-2238	Sequence 2238, Ap	368	6	13.0	254	1	PCT-US01-08631-46154	Sequence 46154, A
296	6	13.0	186	11	US-08-756-693-1	Sequence 1, Appli	369	6	13.0	254	1	US-09-895-828-454	Sequence 454, App
297	6	13.0	186	11	US-08-756-693-9	Sequence 9, Appli	370	6	13.0	255	22	US-60-140-956-1355	Sequence 1355, Ap
298	6	13.0	186	11	US-08-756-693A-9	Sequence 9, Appli	371	6	13.0	255	26	US-60-324-631-1507	Sequence 1507, Ap
299	6	13.0	186	21	US-09-708-427-14540	Sequence 14540, A	372	6	13.0	261	26	US-60-167-217-13019	Sequence 13019, A

373	6	13.0	261	US-60-173-464-10586	Sequence 10586, A	446	6	13.0	297	21	US-09-739-449-11006	Sequence 11006, A
374	6	13.0	261	US-60-191-637-13013	Sequence 10313, A	447	6	13.0	297	21	US-09-789-836-11	Sequence 11, Appl
375	6	13.0	261	US-60-191-681-10235	Sequence 10235, A	448	6	13.0	297	21	US-09-803-110-11006	Sequence 11006, A
376	6	13.0	262	US-09-895-828-457	Sequence 457, Appl	449	6	13.0	299	26	US-60-207-213-317	Sequence 317, Appl
377	6	13.0	265	US-09-708-427-30798	Sequence 30798, A	450	6	13.0	303	16	US-09-270-767-42501	Sequence 42501, A
378	6	13.0	266	US-09-252-991A-19949	Sequence 19949, A	451	6	13.0	305	16	US-09-270-767-38800	Sequence 38800, A
379	6	13.0	268	US-09-248-796-17195	Sequence 17195, A	452	6	13.0	305	16	US-09-270-767-54017	Sequence 54017, A
380	6	13.0	268	US-60-096-409-17195	Sequence 17195, A	453	6	13.0	305	16	US-09-270-849B-189913	Sequence 189913, A
381	6	13.0	268	US-60-191-637-25193	Sequence 25193, A	454	6	13.0	305	16	US-60-208-318-17	Sequence 11, Appl
382	6	13.0	268	US-60-191-681-13832	Sequence 13832, A	455	6	13.0	305	16	US-60-208-965-172	Sequence 172, Appl
383	6	13.0	271	US-09-739-449-8327	Sequence 8327, Appl	456	6	13.0	305	16	US-60-248-505-837	Sequence 837, Appl
384	6	13.0	271	US-09-803-110-8327	Sequence 8327, Appl	457	6	13.0	306	16	US-09-270-767-41820	Sequence 41820, A
385	6	13.0	272	US-60-185-360-245	Sequence 245, Appl	458	6	13.0	307	26	US-60-324-109-30036	Sequence 30036, A
386	6	13.0	273	US-08-761-184-1339	Sequence 1339, Appl	459	6	13.0	308	18	US-09-489-039A-73359	Sequence 73359, Appl
387	6	13.0	273	US-08-761-184-1620	Sequence 1620, Appl	460	6	13.0	312	22	PCT-US01-08631-38763	Sequence 38763, A
388	6	13.0	273	US-08-821-931-1339	Sequence 1339, Appl	461	6	13.0	312	22	US-09-897-516-6721	Sequence 6721, Appl
389	6	13.0	273	US-08-821-931-1620	Sequence 1620, Appl	462	6	13.0	312	22	US-60-215-161-6721	Sequence 6721, Appl
390	6	13.0	273	US-08-993-002A-7630	Sequence 7630, Appl	463	6	13.0	312	26	US-60-324-109-22174	Sequence 22174, A
391	6	13.0	273	US-08-993-002A-7631	Sequence 7631, Appl	464	6	13.0	313	26	US-60-315-676-12	Sequence 12, Appl
392	6	13.0	276	US-60-614-150-35646	Sequence 35646, A	465	6	13.0	317	26	US-09-614-150-35325	Sequence 35325, A
393	6	13.0	276	US-60-191-637-35257	Sequence 35257, A	466	6	13.0	317	26	US-60-173-464-26963	Sequence 26963, A
394	6	13.0	277	1 PCT-US00-35017A-1060	Sequence 1060, Appl	467	6	13.0	317	26	US-60-191-637-34930	Sequence 34930, A
395	6	13.0	277	1 PCT-US97-21976-187	Sequence 187, Appl	468	6	13.0	317	26	US-60-191-681-27412	Sequence 27412, A
396	6	13.0	277	US-09-308-967-187	Sequence 187, Appl	469	6	13.0	321	1	PCT-US02-03987-5771	Sequence 5771, Appl
397	6	13.0	277	US-09-308-967B-187	Sequence 187, Appl	470	6	13.0	321	1	PCT-US02-03987-12694	Sequence 12694, A
398	6	13.0	279	1 PCT-US01-01334-5024	Sequence 5024, Appl	471	6	13.0	321	1	PCT-US02-03987-12721	Sequence 12721, A
399	6	13.0	279	1 PCT-US01-01341-1477	Sequence 1477, Appl	472	6	13.0	321	22	US-09-815-242-5771	Sequence 5771, Appl
400	6	13.0	279	US-09-764-864-1477	Sequence 1477, Appl	473	6	13.0	321	22	US-09-815-242-12694	Sequence 12694, A
401	6	13.0	279	US-09-764-874-5024	Sequence 5024, Appl	474	6	13.0	321	22	US-09-815-242-12721	Sequence 12721, A
402	6	13.0	279	US-10-080-129-1477	Sequence 1477, Appl	475	6	13.0	321	23	US-09-924-359-163	Sequence 163, Appl
403	6	13.0	280	US-09-417-507-41171	Sequence 41171, A	476	6	13.0	321	24	US-10-072-851-5771	Sequence 5771, A
404	6	13.0	280	US-09-643-192-141	Sequence 141, Appl	477	6	13.0	321	24	US-10-072-851-12694	Sequence 12694, A
405	6	13.0	280	US-09-643-192A-141	Sequence 141, Appl	478	6	13.0	321	24	US-10-072-851-12721	Sequence 12721, A
406	6	13.0	282	US-08-448-729-56	Sequence 56, Appl	479	6	13.0	321	26	US-60-242-578-1031	Sequence 1031, Appl
407	6	13.0	282	US-08-449-182-56	Sequence 56, Appl	480	6	13.0	321	26	US-60-253-625-2375	Sequence 2375, Appl
408	6	13.0	282	US-08-449-649-56	Sequence 56, Appl	481	6	13.0	322	1	PCT-US01-05363-6	Sequence 6, Appl
409	6	13.0	282	US-08-993-002A-4789	Sequence 4789, Appl	482	6	13.0	322	1	PCT-US01-05578-6	Sequence 6, Appl
410	6	13.0	282	US-09-005-243-56	Sequence 56, Appl	483	6	13.0	322	21	US-09-789-836-6	Sequence 6, Appl
411	6	13.0	282	US-09-005-858-56	Sequence 56, Appl	484	6	13.0	322	22	US-09-823-038A-33	Sequence 33, Appl
412	6	13.0	282	US-09-005-893-56	Sequence 56, Appl	485	6	13.0	323	18	US-09-438-185-443	Sequence 443, Appl
413	6	13.0	282	US-09-005-907-56	Sequence 56, Appl	486	6	13.0	323	18	US-09-438-185A-443	Sequence 443, Appl
414	6	13.0	282	US-09-005-917-56	Sequence 56, Appl	487	6	13.0	324	26	US-60-315-676-9	Sequence 9, Appl
415	6	13.0	282	US-09-083-838-56	Sequence 56, Appl	488	6	13.0	329	1	PCT-US01-08631-41597	Sequence 41597, A
416	6	13.0	282	US-09-224-683-56	Sequence 56, Appl	489	6	13.0	329	1	PCT-US02-03987-11142	Sequence 11142, A
417	6	13.0	282	US-09-635-249-56	Sequence 56, Appl	490	6	13.0	329	22	US-09-815-242-11142	Sequence 11142, A
418	6	13.0	282	US-09-635-251-56	Sequence 56, Appl	491	6	13.0	329	24	US-10-072-851-11142	Sequence 11142, A
419	6	13.0	282	US-09-643-658A-56	Sequence 56, Appl	492	6	13.0	331	1	PCT-US01-08631-49770	Sequence 49770, A
420	6	13.0	282	US-09-643-659B-56	Sequence 56, Appl	493	6	13.0	331	19	US-09-500-239-14	Sequence 14, Appl
421	6	13.0	282	US-09-758-444-572	Sequence 572, Appl	494	6	13.0	332	20	US-09-611-529-6398	Sequence 6398, Appl
422	6	13.0	283	US-08-761-184-1107	Sequence 1107, Appl	495	6	13.0	332	26	US-60-185-361-840	Sequence 840, Appl
423	6	13.0	283	US-08-821-931-1107	Sequence 1107, Appl	496	6	13.0	334	16	US-09-248-796-20300	Sequence 20300, A
424	6	13.0	283	US-08-993-002A-4788	Sequence 4788, Appl	497	6	13.0	334	16	US-60-096-409-20300	Sequence 20300, A
425	6	13.0	284	US-09-270-767-58734	Sequence 58734, A	498	6	13.0	336	24	US-10-000-897-69	Sequence 69, Appl
426	6	13.0	284	US-09-270-849B-185550	Sequence 185550, A	499	6	13.0	338	21	US-09-739-449-10690	Sequence 10690, A
427	6	13.0	286	1 PCT-US01-08656-10098	Sequence 10098, A	500	6	13.0	338	22	US-09-803-110-10690	Sequence 10690, A
428	6	13.0	288	US-10-015-127-13977	Sequence 13977, A	501	6	13.0	339	1	PCT-US01-08631-57218	Sequence 57218, A
429	6	13.0	289	1 PCT-US01-08631-53361	Sequence 53361, A	502	6	13.0	341	26	US-60-185-361-776	Sequence 776, Appl
430	6	13.0	289	US-09-417-507-28016	Sequence 28016, A	503	6	13.0	342	11	US-08-761-184-1459	Sequence 1459, Appl
431	6	13.0	289	US-09-614-150-9753	Sequence 9753, Appl	504	6	13.0	342	12	US-08-821-931-1459	Sequence 1459, Appl
432	6	13.0	289	US-60-167-217-9791	Sequence 9791, Appl	505	6	13.0	342	13	US-08-993-002A-4790	Sequence 4790, Appl
433	6	13.0	289	US-60-173-464-7897	Sequence 7897, Appl	506	6	13.0	342	13	US-08-993-002A-7045	Sequence 7045, Appl
434	6	13.0	289	US-60-191-637-9785	Sequence 9785, Appl	507	6	13.0	342	20	US-09-614-150-33645	Sequence 33645, A
435	6	13.0	289	US-60-191-681-7631	Sequence 7631, Appl	508	6	13.0	342	22	US-09-815-108-20	Sequence 20, Appl
436	6	13.0	292	US-09-051-755-10	Sequence 10, Appl	509	6	13.0	342	26	US-60-173-464-22113	Sequence 22113, A
437	6	13.0	294	US-09-902-540-12714	Sequence 12714, A	510	6	13.0	342	26	US-60-191-637-32223	Sequence 32223, A
438	6	13.0	295	US-09-739-449-9863	Sequence 9863, Appl	511	6	13.0	342	26	US-60-191-681-26158	Sequence 26158, A
439	6	13.0	295	US-09-603-110-9863	Sequence 9863, Appl	512	6	13.0	344	1	PCT-US01-08631-33558	Sequence 33558, A
440	6	13.0	296	1 PCT-US01-05363-9	Sequence 9, Appl	513	6	13.0	344	21	US-09-712-363-165	Sequence 165, Appl
441	6	13.0	296	1 PCT-US01-05578-9	Sequence 9, Appl	514	6	13.0	344	23	US-09-902-540-11424	Sequence 11424, A
442	6	13.0	296	US-09-270-767-44827	Sequence 44827, A	515	6	13.0	345	1	PCT-US01-05363-5	Sequence 5, Appl
443	6	13.0	296	US-09-789-836-9	Sequence 9, Appl	516	6	13.0	345	1	PCT-US01-05578-5	Sequence 5, Appl
444	6	13.0	297	1 PCT-US01-05363-18	Sequence 18, Appl	517	6	13.0	345	1	PCT-US01-10661-10	Sequence 10, Appl
445	6	13.0	297	1 PCT-US01-05578-11	Sequence 11, Appl	518	6	13.0	345	1	PCT-US02-02781-183	Sequence 183, Appl

519	6	13.0	345	21	US-09-789-836-5	Sequence 5, Appli	592	6	13.0	392	20	US-09-619-049-108	Sequence 108, App
520	6	13.0	345	24	US-10-060-036-183	Sequence 183, App	593	6	13.0	392	26	US-60-167-217-1530	Sequence 1530, App
521	6	13.0	350	1	PCT-US00-18198-103	Sequence 103, App	594	6	13.0	392	26	US-60-171-627-144	Sequence 144, App
522	6	13.0	350	17	PCT-US02-03987-15264	Sequence 15264, A	595	6	13.0	392	26	US-60-173-464-1242	Sequence 1242, App
523	6	13.0	350	17	US-09-345-464-103	Sequence 103, App	596	6	13.0	392	26	US-60-191-637-1506	Sequence 1506, App
524	6	13.0	350	20	US-09-606-317-103	Sequence 103, App	597	6	13.0	392	26	US-60-191-681-1194	Sequence 1194, App
525	6	13.0	350	24	US-10-032-585-7320	Sequence 7320, App	598	6	13.0	395	22	US-09-897-516-5816	Sequence 5816, App
526	6	13.0	350	24	US-10-072-851-15264	Sequence 15264, A	599	6	13.0	395	26	US-60-215-161-5816	Sequence 5816, App
527	6	13.0	350	26	US-60-189-998-3	Sequence 3, Appli	600	6	13.0	396	18	US-09-470-934A-49	Sequence 49, Appli
528	6	13.0	350	26	US-60-259-128-4741	Sequence 4741, App	601	6	13.0	397	26	US-60-173-903-2	Sequence 2, Appli
529	6	13.0	350	26	US-08-487-032A-7320	Sequence 7320, App	602	6	13.0	398	1	PCT-US01-08631-41200	Sequence 41200, A
530	6	13.0	351	8	US-08-561-469A-640	Sequence 640, App	603	6	13.0	404	1	PCT-US00-26524B-6307	Sequence 6307, App
531	6	13.0	351	11	US-08-761-184-983	Sequence 983, App	604	6	13.0	404	20	US-09-614-150-21411	Sequence 21411, A
532	6	13.0	351	12	US-08-821-931-983	Sequence 983, App	605	6	13.0	404	26	US-60-167-217-21620	Sequence 21620, A
533	6	13.0	351	13	US-08-993-002A-4795	Sequence 4795, App	606	6	13.0	404	26	US-60-173-464-17585	Sequence 17585, A
534	6	13.0	352	20	US-09-699-146-244	Sequence 244, App	607	6	13.0	404	26	US-60-191-637-21473	Sequence 21473, A
535	6	13.0	354	11	US-08-761-184-1560	Sequence 1560, App	608	6	13.0	406	1	US-60-191-681-16925	Sequence 16925, A
536	6	13.0	354	12	US-08-821-931-1560	Sequence 1560, App	609	6	13.0	406	1	PCT-US01-14827-11742	Sequence 11742, A
537	6	13.0	354	13	US-08-993-002A-4796	Sequence 4796, App	610	6	13.0	410	1	PCT-US00-30036A-84	Sequence 84, Appli
538	6	13.0	355	11	US-08-761-184-1161	Sequence 1161, App	611	6	13.0	410	1	PCT-US01-11988-710	Sequence 710, App
539	6	13.0	355	12	US-08-821-931-1161	Sequence 1161, App	612	6	13.0	410	1	PCT-US02-08123-1365	Sequence 1365, App
540	6	13.0	355	13	US-08-993-002A-4797	Sequence 4797, App	613	6	13.0	410	1	PCT-US02-08277-890	Sequence 890, App
541	6	13.0	356	18	US-09-417-507-28013	Sequence 28013, A	614	6	13.0	410	1	PCT-US02-08278-1312	Sequence 1312, App
542	6	13.0	356	18	US-09-488-725A-2917	Sequence 2917, App	615	6	13.0	410	22	US-09-833-245-710	Sequence 710, App
543	6	13.0	356	18	US-09-488-725A-2917	Sequence 2917, App	616	6	13.0	412	9	US-08-553-125-12	Sequence 12, Appli
544	6	13.0	357	16	US-09-252-991A-25036	Sequence 25036, A	617	6	13.0	412	9	US-08-553-125A-12	Sequence 12, Appli
545	6	13.0	358	23	US-09-933-480-3	Sequence 3, Appli	618	6	13.0	412	11	US-08-798-095-12	Sequence 12, Appli
546	6	13.0	359	20	US-09-614-150-39777	Sequence 39777, A	619	6	13.0	412	11	US-08-798-095A-12	Sequence 12, Appli
547	6	13.0	359	26	US-60-146-315-913	Sequence 913, App	620	6	13.0	412	11	US-08-798-095-12	Sequence 12, Appli
548	6	13.0	359	26	US-60-173-464-29633	Sequence 29633, A	621	6	13.0	412	13	US-08-913-095-12	Sequence 12, Appli
549	6	13.0	359	26	US-60-173-464-29633	Sequence 29633, A	622	6	13.0	412	13	US-08-913-095-12	Sequence 12, Appli
550	6	13.0	359	26	US-60-191-637-39409	Sequence 39409, A	623	6	13.0	412	13	US-08-948-538A-41	Sequence 41, Appli
551	6	13.0	359	26	US-60-191-681-30457	Sequence 30457, A	624	6	13.0	412	16	US-09-215-450-25	Sequence 25, Appli
552	6	13.0	360	1	PCT-US01-08631-45056	Sequence 45056, A	625	6	13.0	412	16	US-09-219-441-12	Sequence 12, Appli
553	6	13.0	360	15	US-09-198-452A-342	Sequence 342, App	626	6	13.0	412	17	US-09-380-159-6	Sequence 6, Appli
554	6	13.0	360	17	US-09-376-770-6	Sequence 6, Appli	627	6	13.0	412	22	US-09-882-636-6	Sequence 6, Appli
555	6	13.0	360	18	US-09-488-725A-6489	Sequence 6489, App	628	6	13.0	412	22	US-09-953-956-12	Sequence 12, Appli
556	6	13.0	360	20	US-09-614-150-39240	Sequence 39240, A	629	6	13.0	412	22	US-09-897-516-5849	Sequence 5849, App
557	6	13.0	360	26	US-60-191-637-38866	Sequence 38866, A	630	6	13.0	412	22	US-60-167-217-3850	Sequence 3850, App
558	6	13.0	360	26	US-60-191-681-30124	Sequence 30124, A	631	6	13.0	413	26	US-60-173-464-3176	Sequence 3176, App
559	6	13.0	365	1	PCT-US00-35017A-1026	Sequence 1026, App	632	6	13.0	413	26	US-60-215-161-6849	Sequence 6849, App
560	6	13.0	366	18	US-09-438-185-324	Sequence 324, App	633	6	13.0	413	26	US-60-245-228-383	Sequence 383, App
561	6	13.0	366	18	US-09-438-185A-334	Sequence 334, App	634	6	13.0	413	26	US-60-213-177-637	Sequence 637, App
562	6	13.0	367	15	US-09-134-001C-3678	Sequence 3678, App	635	6	13.0	414	26	US-60-185-361-534	Sequence 534, App
563	6	13.0	368	1	PCT-US01-08631-45056	Sequence 108, App	636	6	13.0	422	21	US-09-733-089-15567	Sequence 15567, A
564	6	13.0	368	1	PCT-US01-11988-707	Sequence 707, App	637	6	13.0	422	21	US-09-815-660-15567	Sequence 15567, A
565	6	13.0	368	1	PCT-US02-08123-1751	Sequence 1751, App	638	6	13.0	422	21	US-09-815-660-15567	Sequence 15567, A
566	6	13.0	368	1	PCT-US02-08277-1146	Sequence 1146, App	639	6	13.0	423	20	US-09-614-150-14088	Sequence 14088, A
567	6	13.0	368	1	PCT-US02-08278-1683	Sequence 1683, App	640	6	13.0	423	26	US-60-191-637-14129	Sequence 14129, A
568	6	13.0	368	22	US-09-833-245-707	Sequence 707, App	641	6	13.0	423	16	US-09-252-691-9032	Sequence 9032, App
569	6	13.0	370	15	US-09-107-433-3702	Sequence 3702, App	642	6	13.0	435	16	US-09-252-691-9032	Sequence 9032, App
570	6	13.0	370	19	US-09-583-110-3405	Sequence 3405, App	643	6	13.0	435	16	US-60-167-216-74	Sequence 74, Appli
571	6	13.0	371	12	US-08-827-356-3712	Sequence 3712, App	644	6	13.0	435	26	US-60-173-464-2285	Sequence 2285, App
572	6	13.0	371	18	US-09-488-725A-5360	Sequence 5360, App	645	6	13.0	438	1	PCT-US01-43782-2	Sequence 2, Appli
573	6	13.0	371	20	US-09-611-529-4501	Sequence 4501, App	646	6	13.0	438	24	US-10-001-848-2	Sequence 2, Appli
574	6	13.0	373	22	US-09-823-038A-60	Sequence 60, Appli	647	6	13.0	439	22	US-09-823-038A-32	Sequence 32, Appli
575	6	13.0	375	21	US-09-708-427-10518	Sequence 10518, A	648	6	13.0	442	1	PCT-US02-03987-11406	Sequence 11406, A
576	6	13.0	377	1	PCT-US01-26213-2	Sequence 2, Appli	649	6	13.0	442	1	PCT-US98-06371-1014	Sequence 1014, App
577	6	13.0	377	13	US-08-988-671A-31	Sequence 31, Appli	650	6	13.0	442	13	US-08-902-615A-280	Sequence 280, App
578	6	13.0	377	16	US-09-270-767-46304	Sequence 46304, A	651	6	13.0	442	22	US-08-993-002A-7046	Sequence 7046, App
579	6	13.0	377	18	US-09-488-725A-1978	Sequence 1978, A	652	6	13.0	442	22	US-09-815-242-11406	Sequence 11406, A
580	6	13.0	377	20	US-09-611-526-2593	Sequence 2593, App	653	6	13.0	442	24	US-10-072-851-11406	Sequence 11406, A
581	6	13.0	377	20	US-09-680-985-2	Sequence 2, Appli	654	6	13.0	444	16	US-09-270-767-33740	Sequence 33740, A
582	6	13.0	377	23	US-09-935-038-2	Sequence 2, Appli	655	6	13.0	444	16	US-09-270-767-48957	Sequence 48957, A
583	6	13.0	377	23	US-09-935-038A-31	Sequence 31, Appli	656	6	13.0	444	16	US-09-270-767-48957	Sequence 48957, A
584	6	13.0	377	24	US-10-094-474-2	Sequence 2, Appli	657	6	13.0	446	1	PCT-US99-05606-51	Sequence 51, Appli
585	6	13.0	377	26	US-60-213-845-443	Sequence 443, App	658	6	13.0	446	20	US-09-631-275-51	Sequence 51, Appli
586	6	13.0	378	22	US-09-833-750-256	Sequence 256, App	659	6	13.0	446	20	US-09-657-474-51	Sequence 51, Appli
587	6	13.0	385	18	US-09-488-725A-5550	Sequence 5550, App	660	6	13.0	447	21	US-09-708-427-1527	Sequence 1527, App
588	6	13.0	385	19	US-09-570-581A-885	Sequence 885, App	661	6	13.0	447	26	US-60-208-129-190	Sequence 190, App
589	6	13.0	388	26	US-60-315-676-2034	Sequence 2034, App	662	6	13.0	447	26	US-60-208-973-433	Sequence 233, App
590	6	13.0	390	26	US-60-221-680-5	Sequence 5, Appli	663	6	13.0	448	22	US-09-815-108-6	Sequence 6, Appli
591	6	13.0	392	20	US-09-614-150-1503	Sequence 1503, App	664	6	13.0	449	21	US-09-708-427-2162	Sequence 2162, App

665	6	13.0	450	16	US-09-270-767-46295	Sequence 46295, A	738	6	13.0	504	21	US-09-709-238-119	Sequence 119, App
666	6	13.0	451	1	PCT-US01-14827-10545	Sequence 10545, A	739	6	13.0	504	21	US-09-738-386-2	Sequence 2, Appl
667	6	13.0	451	17	US-09-328-352-7659	Sequence 7659, Ap	740	6	13.0	504	21	US-09-756-753-94	Sequence 94, Appl
668	6	13.0	452	15	US-09-126-642-8	Sequence 8, Appl	741	6	13.0	504	22	US-09-796-753-108	Sequence 108, App
669	6	13.0	452	18	US-09-463-713C-8	Sequence 8, Appl	742	6	13.0	504	22	US-09-815-108-8	Sequence 8, Appl
670	6	13.0	452	22	US-09-865-415-6	Sequence 6, Appl	743	6	13.0	504	22	US-09-815-108-15	Sequence 15, Appl
671	6	13.0	455	20	US-09-602-874C-54	Sequence 54, Appl	744	6	13.0	504	22	US-09-815-108-17	Sequence 17, Appl
672	6	13.0	455	20	US-09-620-394B-6631	Sequence 6631, Ap	745	6	13.0	504	22	US-09-815-108-19	Sequence 19, Appl
673	6	13.0	455	21	US-09-738-626-4477	Sequence 12, Appl	746	6	13.0	504	23	US-09-933-480-2	Sequence 2, Appl
674	6	13.0	456	19	US-09-557-262-12	Sequence 12, Appl	747	6	13.0	504	23	US-09-941-992-119	Sequence 119, App
675	6	13.0	463	1	PCT-US97-05223-1068	Sequence 1068, Ap	748	6	13.0	504	23	US-09-989-279-119	Sequence 119, App
676	6	13.0	463	12	US-08-824-132-1068	Sequence 1068, Ap	749	6	13.0	504	23	US-09-989-293A-119	Sequence 119, App
677	6	13.0	463	13	US-08-993-002A-5000	Sequence 5000, Ap	750	6	13.0	504	23	US-09-989-721-119	Sequence 119, App
678	6	13.0	465	1	PCT-US01-08631-33517	Sequence 33517, A	751	6	13.0	504	23	US-09-989-722-119	Sequence 119, App
679	6	13.0	466	1	PCT-US99-05606-64	Sequence 64, Appl	752	6	13.0	504	23	US-09-989-723-119	Sequence 119, App
680	6	13.0	466	20	US-09-620-394B-6630	Sequence 6630, Ap	753	6	13.0	504	23	US-09-989-724-119	Sequence 119, App
681	6	13.0	466	20	US-09-631-275-64	Sequence 64, Appl	754	6	13.0	504	23	US-09-989-725-119	Sequence 119, App
682	6	13.0	466	20	US-09-657-474-64	Sequence 64, Appl	755	6	13.0	504	23	US-09-989-726-119	Sequence 119, App
683	6	13.0	467	18	US-09-428-944-2754	Sequence 2754, Ap	756	6	13.0	504	23	US-09-989-727-119	Sequence 119, App
684	6	13.0	468	21	US-09-708-427-1526	Sequence 1526, Ap	757	6	13.0	504	23	US-09-989-728-119	Sequence 119, App
685	6	13.0	472	22	US-09-815-108-5	Sequence 5, Appl	758	6	13.0	504	23	US-09-989-729A-119	Sequence 119, App
686	6	13.0	473	21	US-09-708-427-31892	Sequence 31892, A	759	6	13.0	504	23	US-09-989-730-119	Sequence 119, App
687	6	13.0	477	1	PCT-US99-05606-4	Sequence 4, Appl	760	6	13.0	504	23	US-09-989-731-119	Sequence 119, App
688	6	13.0	477	16	US-09-236-134-4	Sequence 4, Appl	761	6	13.0	504	23	US-09-989-732-119	Sequence 119, App
689	6	13.0	477	16	US-09-236-134A-4	Sequence 4, Appl	762	6	13.0	504	23	US-09-989-734-119	Sequence 119, App
690	6	13.0	477	16	US-09-236-134B-4	Sequence 4, Appl	763	6	13.0	504	23	US-09-989-735-119	Sequence 119, App
691	6	13.0	477	16	US-09-236-134C-4	Sequence 4, Appl	764	6	13.0	504	23	US-09-989-862-119	Sequence 119, App
692	6	13.0	477	20	US-09-631-275-4	Sequence 4, Appl	765	6	13.0	504	23	US-09-990-427-119	Sequence 119, App
693	6	13.0	477	20	US-09-657-474-4	Sequence 4, Appl	766	6	13.0	504	23	US-09-990-436-119	Sequence 119, App
694	6	13.0	477	21	US-09-722-544A-4	Sequence 4, Appl	767	6	13.0	504	23	US-09-990-437-119	Sequence 119, App
695	6	13.0	478	23	US-09-971-309-19	Sequence 19, Appl	768	6	13.0	504	23	US-09-990-438-119	Sequence 119, App
696	6	13.0	478	20	US-09-614-150-3780	Sequence 3780, Ap	769	6	13.0	504	23	US-09-990-439-119	Sequence 119, App
697	6	13.0	478	26	US-60-191-637-3792	Sequence 3792, Ap	770	6	13.0	504	23	US-09-990-440-119	Sequence 119, App
698	6	13.0	478	26	US-60-191-681-3048	Sequence 3048, Ap	771	6	13.0	504	23	US-09-990-441-119	Sequence 119, App
699	6	13.0	479	19	US-09-583-110-4687	Sequence 4687, Ap	772	6	13.0	504	23	US-09-990-442-119	Sequence 119, App
700	6	13.0	480	1	PCT-US00-18198-102	Sequence 102, App	773	6	13.0	504	23	US-09-990-443-119	Sequence 119, App
701	6	13.0	480	15	US-09-107-433-3418	Sequence 3418, Ap	774	6	13.0	504	23	US-09-990-444-119	Sequence 119, App
702	6	13.0	480	17	US-09-345-464-102	Sequence 102, App	775	6	13.0	504	23	US-09-990-456-119	Sequence 119, App
703	6	13.0	480	20	US-09-606-317-102	Sequence 102, App	776	6	13.0	504	23	US-09-990-562-119	Sequence 119, App
704	6	13.0	483	16	US-09-252-991A-20139	Sequence 20139, A	777	6	13.0	504	23	US-09-990-711-119	Sequence 119, App
705	6	13.0	484	17	US-09-389-956-12	Sequence 12, Appl	778	6	13.0	504	23	US-09-990-726-119	Sequence 119, App
706	6	13.0	485	16	US-09-270-767-43385	Sequence 43385, A	779	6	13.0	504	23	US-09-991-073-119	Sequence 119, App
707	6	13.0	488	26	US-60-213-178-598	Sequence 598, App	780	6	13.0	504	23	US-09-991-157-119	Sequence 119, App
708	6	13.0	489	18	US-09-417-507-29712	Sequence 29712, A	781	6	13.0	504	23	US-09-991-163-119	Sequence 119, App
709	6	13.0	489	20	US-09-620-394B-6629	Sequence 6629, Ap	782	6	13.0	504	23	US-09-991-172-119	Sequence 119, App
710	6	13.0	491	13	US-08-993-002A-6529	Sequence 6529, Ap	783	6	13.0	504	23	US-09-991-181-119	Sequence 119, App
711	6	13.0	493	23	US-09-918-951-5	Sequence 5, Appl	784	6	13.0	504	23	US-09-991-854-119	Sequence 119, App
712	6	13.0	494	18	US-09-417-507-42209	Sequence 42209, A	785	6	13.0	504	23	US-09-992-521-119	Sequence 119, App
713	6	13.0	494	20	US-09-631-275-133	Sequence 133, App	786	6	13.0	504	23	US-09-992-598-119	Sequence 119, App
714	6	13.0	495	1	PCT-US99-05606-2	Sequence 2, Appl	787	6	13.0	504	23	US-09-993-469-119	Sequence 119, App
715	6	13.0	495	16	US-09-236-134-2	Sequence 2, Appl	788	6	13.0	504	23	US-09-993-583-119	Sequence 119, App
716	6	13.0	495	16	US-09-236-134A-2	Sequence 2, Appl	789	6	13.0	504	23	US-09-993-604-119	Sequence 119, App
717	6	13.0	495	16	US-09-236-134B-2	Sequence 2, Appl	790	6	13.0	504	23	US-09-993-667-119	Sequence 119, App
718	6	13.0	495	16	US-09-236-134C-2	Sequence 2, Appl	791	6	13.0	504	23	US-09-993-687-119	Sequence 119, App
719	6	13.0	495	20	US-09-657-474-2	Sequence 2, Appl	792	6	13.0	504	23	US-09-993-748-119	Sequence 119, App
720	6	13.0	495	21	US-09-722-544A-2	Sequence 2, Appl	793	6	13.0	504	23	US-09-994-054-119	Sequence 119, App
721	6	13.0	495	20	US-09-722-544A-2	Sequence 2, Appl	794	6	13.0	504	23	US-09-996-243-119	Sequence 119, App
722	6	13.0	496	16	US-09-201-228A-423	Sequence 423, App	795	6	13.0	504	23	US-09-997-333-119	Sequence 119, App
723	6	13.0	496	22	US-09-808-483-14	Sequence 14, Appl	796	6	13.0	504	23	US-09-997-349-119	Sequence 119, App
724	6	13.0	499	26	US-60-213-177-791	Sequence 791, App	797	6	13.0	504	23	US-09-997-384-119	Sequence 119, App
725	6	13.0	501	20	US-09-614-150-21603	Sequence 21603, A	798	6	13.0	504	23	US-09-997-428-119	Sequence 119, App
726	6	13.0	501	26	US-60-167-217-21814	Sequence 21814, A	799	6	13.0	504	23	US-09-997-440-119	Sequence 119, App
727	6	13.0	501	26	US-60-173-464-17759	Sequence 17759, A	800	6	13.0	504	23	US-09-997-514-119	Sequence 119, App
728	6	13.0	501	26	US-60-191-637-21668	Sequence 21668, A	801	6	13.0	504	23	US-09-997-529-119	Sequence 119, App
729	6	13.0	501	26	US-60-191-681-17093	Sequence 17093, A	802	6	13.0	504	23	US-09-997-542-119	Sequence 119, App
730	6	13.0	502	15	US-09-134-000-6324	Sequence 6324, Ap	803	6	13.0	504	23	US-09-997-559-119	Sequence 119, App
731	6	13.0	504	1	PCT-US01-18198-5	Sequence 5, Appl	804	6	13.0	504	23	US-09-997-573-119	Sequence 119, App
732	6	13.0	504	1	PCT-US01-43782-3	Sequence 3, Appl	805	6	13.0	504	23	US-09-997-585-119	Sequence 119, App
733	6	13.0	504	1	PCT-US99-08526-2	Sequence 2, Appl	806	6	13.0	504	23	US-09-997-601-119	Sequence 119, App
734	6	13.0	504	16	US-09-293-182-2	Sequence 2, Appl	807	6	13.0	504	23	US-09-997-614-119	Sequence 119, App
735	6	13.0	504	17	US-09-345-464-5	Sequence 5, Appl	808	6	13.0	504	23	US-09-997-628-119	Sequence 119, App
736	6	13.0	504	17	US-09-388-778-2	Sequence 2, Appl	809	6	13.0	504	23	US-09-997-641-119	Sequence 119, App
737	6	13.0	504	20	US-09-606-317-5	Sequence 5, Appl	810	6	13.0	504	23	US-09-997-653-119	Sequence 119, App

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 11:02:02 : Search time 17.8507 Seconds
(without alignments)
247.615 Million cell updates/sec

Title: US-09-833-017b-2
Perfect score: 46
Sequence: 1 MKKTLNKFKEIKTDELE.....GSLSTPFRLEFRSFTQALGK 46

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR71.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	19.6	971	2 D70128	conserved hypoteth
2	7	15.2	88	2 G90517	hypothetical prote
3	7	15.2	230	2 S46737	hypothetical prote
4	7	15.2	240	2 F83488	probable Atp-bind
5	7	15.2	240	2 H95873	probable amino aci
6	7	15.2	240	2 G98155	probable Atp-bind
7	7	15.2	240	2 AE3132	hypothetical prote
8	7	15.2	253	2 B70483	ABC transporter -
9	7	15.2	287	2 S70366	stem cell factor s
10	7	15.2	370	2 T48633	hypothetical prote
11	7	15.2	411	1 I55604	platelet glycoprot
12	7	15.2	437	2 AC1112	pns beta-glucoside
13	7	15.2	437	2 AG1473	pns beta-glucoside
14	7	15.2	472	2 AD2284	hypothetical prote
15	7	15.2	494	2 B96534	hypothetical prote
16	7	15.2	558	2 A40223	transcription acti
17	7	15.2	565	2 A10479	probable membrane
18	7	15.2	579	2 A42617	66k rel-related pr
19	7	15.2	591	2 S51303	hypothetical prote
20	7	15.2	618	2 D86364	hypothetical prote
21	7	15.2	788	2 T07667	UDPglucose--glycoq
22	7	15.2	808	2 D81924	hypothetical prote
23	7	15.2	808	2 D81924	hypothetical prote
24	7	15.2	1183	2 F90559	conserved hypoteth
25	7	15.2	1358	2 T22695	hypothetical prote
26	7	15.2	1804	2 AF2250	serine/threonine k
27	7	15.2	3972	2 S75251	hypothetical prote
28	6	13.0	42	2 A31918	cathepsin D (EC 3.
29	6	13.0	43	2 A97416	hypothetical prote

30	6	13.0	48	2	H82592	hypothetical prote
31	6	13.0	69	2	T18094	hypothetical prote
32	6	13.0	88	2	QXDO	26S proteasome reg
33	6	13.0	97	2	H96743	probable DnaI prot
34	6	13.0	102	2	G71137	hypothetical prote
35	6	13.0	104	2	B81045	hypothetical prote
36	6	13.0	106	2	AE1842	multidrug exporter
37	6	13.0	106	2	D97574	hypothetical prote
38	6	13.0	106	2	AD2795	conserved hypoteth
39	6	13.0	110	2	S73121	cytochrome c553 -
40	6	13.0	114	1	M3HWE	ig heavy chain V-I
41	6	13.0	122	1	AIHUTR	ig heavy chain V-I
42	6	13.0	122	2	B75387	hypothetical prote
43	6	13.0	127	2	A82324	hypothetical prote
44	6	13.0	128	2	AE1330	hypothetical prote
45	6	13.0	133	2	H87681	hypothetical prote
46	6	13.0	134	2	T17655	hypothetical prote
47	6	13.0	141	2	H86609	kinase [imported]
48	6	13.0	152	2	AD1071	conserved hypoteth
49	6	13.0	154	2	G86699	hypothetical prote
50	6	13.0	155	2	B70370	hypothetical prote
51	6	13.0	159	2	AI0772	probable exported
52	6	13.0	181	2	S78572	EST3 protein - yea
53	6	13.0	184	2	AD0207	probable exported
54	6	13.0	186	2	A56956	GTP-binding protei
55	6	13.0	197	2	A44363	voltage-gated dihy
56	6	13.0	202	2	H81731	CDP-diacylglycerol
57	6	13.0	202	2	JX0228	multicatalytic end
58	6	13.0	202	2	I49121	multicatalytic end
59	6	13.0	203	2	C95120	hypothetical prote
60	6	13.0	203	2	G97989	hypothetical prote
61	6	13.0	205	2	S17522	multicatalytic end
62	6	13.0	209	2	A56632	proteasome LMP2.s
63	6	13.0	213	2	JE0101	proteasome subunit
64	6	13.0	213	2	E71511	hypothetical prote
65	6	13.0	214	2	G70348	conserved hypoteth
66	6	13.0	216	1	XXSACC	chloramphenicol O-
67	6	13.0	219	2	JX0231	proteasome ring12
68	6	13.0	219	2	JC2019	multicatalytic end
69	6	13.0	219	2	S27332	proteasome chain L
70	6	13.0	220	2	B64681	nicotinamide monon
71	6	13.0	220	2	B71832	nicotinamide monon
72	6	13.0	220	2	E98293	hypothetical prote
73	6	13.0	220	2	AD2990	conserved hypoteth
74	6	13.0	224	2	T36291	probable glycine-r
75	6	13.0	226	2	T40487	proteasome compone
76	6	13.0	227	2	S66482	transcription regu
77	6	13.0	229	2	S66342	ribonuclease II (E
78	6	13.0	232	2	A99554	hypothetical prote
79	6	13.0	232	2	T04497	multicatalytic end
80	6	13.0	234	2	T03985	multicatalytic end
81	6	13.0	234	2	E84177	hypothetical prote
82	6	13.0	235	2	C82155	probable C-factor
83	6	13.0	236	2	T50209	translin-like prot
84	6	13.0	237	2	A64352	hypothetical prote
85	6	13.0	239	2	B54589	proteasome subunit
86	6	13.0	243	2	D72098	serine esterase, p
87	6	13.0	243	2	G86524	lysophospholipase
88	6	13.0	243	2	C95295	probable ABC trans
89	6	13.0	256	2	F87707	ABC transporter, A
90	6	13.0	257	2	H75199	methythioadenosin
91	6	13.0	260	2	C71233	probable 5'-methyl
92	6	13.0	262	2	G82901	conserved hypoteth
93	6	13.0	265	2	G84743	hypothetical prote
94	6	13.0	270	2	G90404	5'-methylthiadenos
95	6	13.0	271	2	G84030	hema concentration
96	6	13.0	273	2	F64590	diaminopimelate ep
97	6	13.0	273	2	G71923	diaminopimelate ep
98	6	13.0	274	2	A83981	hypothetical prote
99	6	13.0	274	2	A11729	gp17 (Bacteriophag
100	6	13.0	276	2	D41044	octopine-binding p
101	6	13.0	277	2	C70351	purine nucleoside
102	6	13.0	278	2	B97069	anaerobic Cobalt c

103	6	13.0	283	2	C23660	calcium channel pr	176	6	13.0	407	1	KHRTD	cathepsin D (EC 3.
104	6	13.0	285	2	T29047	hypothetical prote	177	6	13.0	410	1	KHMSD	cathepsin D (EC 3.
105	6	13.0	285	2	E83795	ABC transporter (A	178	6	13.0	412	1	KHMSD	probable NAD(P)-de
106	6	13.0	287	2	JN0637	stem cell factor p	179	6	13.0	416	2	AC1143	probable NAD(P)-de
107	6	13.0	287	2	D97241	phosphohydrolase f	180	6	13.0	416	2	AF1501	UDP-N-acetylglucos
108	6	13.0	288	2	D82578	conserved hypotet	181	6	13.0	418	2	B81359	sulfite reductase
109	6	13.0	289	2	T28922	hypothetical prote	182	6	13.0	418	2	G69302	probable protein (
110	6	13.0	294	2	D20256	ribose phosphate p	183	6	13.0	421	2	C95880	hypothetical prote
111	6	13.0	294	2	A23660	calcium channel pr	184	6	13.0	422	2	T33374	histidyl-tRNA synt
112	6	13.0	294	2	B23660	calcium channel pr	185	6	13.0	427	2	F75406	hypothetical prote
113	6	13.0	295	2	S05006	calcium channel al	186	6	13.0	428	2	T31746	26S Proteasome reg
114	6	13.0	295	2	F98205	probable transcrip	187	6	13.0	429	2	S59773	ntra-like serine p
115	6	13.0	295	2	AC3081	transcription regu	188	6	13.0	433	2	H97199	aspartate oxidase
116	6	13.0	296	2	A36366	enhancer-binding p	189	6	13.0	434	2	E97026	histidine--tRNA li
117	6	13.0	297	2	A35914	transcription fact	190	6	13.0	442	1	F64668	histidine--tRNA li
118	6	13.0	297	2	E97387	hypothetical prote	191	6	13.0	442	2	D71847	probable disease r
119	6	13.0	297	2	AF2605	oxidoreductase Atu	192	6	13.0	447	2	G84687	hypothetical prote
120	6	13.0	298	2	H83863	modulation of CheA	193	6	13.0	457	2	H85013	probable glycosyl
121	6	13.0	300	2	E83388	probable transcrip	194	6	13.0	460	2	AG0965	1,2-diacylglycerol
122	6	13.0	303	2	H82422	hypothetical prote	195	6	13.0	468	2	T52269	probable sterol o-
123	6	13.0	303	2	A30443	dihydrodipicolinat	196	6	13.0	472	2	T41684	alkaline metallopr
124	6	13.0	304	2	C71163	probable oligopept	197	6	13.0	476	2	A41463	hypothetical prote
125	6	13.0	305	2	B82873	conserved hypotet	198	6	13.0	478	2	H71224	alkaline metallopr
126	6	13.0	306	2	B75142	oligopeptide trans	199	6	13.0	479	1	S26699	potassium uptake p
127	6	13.0	306	2	T08504	trbG protein - Ent	200	6	13.0	479	2	E95055	probable ss-DNA-sp
128	6	13.0	309	2	G75068	abc transporter PA	201	6	13.0	481	2	G75192	hypothetical prote
129	6	13.0	309	2	JC4282	gamma-hemolysin II	202	6	13.0	484	2	T50909	RI01 protein - yea
130	6	13.0	309	2	D90043	gamma-hemolysin ch	203	6	13.0	484	2	H70573	hypothetical prote
131	6	13.0	311	2	D64184	oligopeptide trans	204	6	13.0	493	2	S61006	hypothetical prote
132	6	13.0	312	2	T24625	hypothetical prote	205	6	13.0	494	2	S48769	hypothetical prote
133	6	13.0	314	2	T45010	probable site-spec	206	6	13.0	497	2	D59853	altronate hydrolas
134	6	13.0	314	2	F84987	methionyl-tRNA for	207	6	13.0	499	1	DELVW4	NADH dehydrogenase
135	6	13.0	316	2	G20277	ct007 hypothetical	208	6	13.0	499	2	S39113	alpha-L-arabinofur
136	6	13.0	316	2	G86545	CT007 hypothetical	209	6	13.0	506	2	C64377	hypothetical prote
137	6	13.0	317	1	B41671	iron transport pro	210	6	13.0	508	2	T22440	hypothetical prote
138	6	13.0	321	2	S49271	hlga-like protein	211	6	13.0	518	2	H69649	2-isopropylmalate
139	6	13.0	323	2	H90119	hypothetical prote	212	6	13.0	519	2	T24243	hypothetical prote
140	6	13.0	326	2	S76441	hypothetical prote	213	6	13.0	521	2	S08435	gag polyprotein -
141	6	13.0	329	1	D71316	conserved hypotet	214	6	13.0	530	2	T43089	transfer complex p
142	6	13.0	329	2	I64160	hypothetical prote	215	6	13.0	538	2	I51560	ribonucleoprotein
143	6	13.0	330	2	C75065	hypothetical prote	216	6	13.0	550	1	A42679	protein-tyrosine-p
144	6	13.0	338	2	AC7556	hypothetical prote	217	6	13.0	560	2	D71913	lytic murein trans
145	6	13.0	338	2	AC2776	pyridine nucleotid	218	6	13.0	560	2	E64600	soluble lytic mure
146	6	13.0	340	2	T42995	flagellar motor sw	219	6	13.0	561	1	E64980	hypothetical 62.1
147	6	13.0	342	2	A71957	flagellar p-ring p	220	6	13.0	561	2	A98996	probable 2-compone
148	6	13.0	342	2	F64550	flagellar basal-bo	221	6	13.0	561	2	F85841	probable two-compo
149	6	13.0	344	1	KHPGD	cathepsin D (EC 3.	222	6	13.0	561	2	AE0777	hypothetical prote
150	6	13.0	344	2	D70576	probable fructose	223	6	13.0	567	2	E84658	hypothetical prote
151	6	13.0	345	2	S12788	transcription fact	224	6	13.0	569	2	G81400	proline-tRNA liga
152	6	13.0	349	2	S64003	hypothetical prote	225	6	13.0	569	2	S11035	chaperonin hsp60
153	6	13.0	354	1	G64648	flagellar motor sw	226	6	13.0	571	2	C83527	prolyl-tRNA synthe
154	6	13.0	354	2	G71938	flagellar motor sw	227	6	13.0	572	2	C42679	protein-tyrosine-p
155	6	13.0	360	2	B86531	Yop translocation	228	6	13.0	573	1	A32800	chaperonin GroEL p
156	6	13.0	360	2	C72091	type III secretion	229	6	13.0	573	1	HHR760	chaperonin groEL p
157	6	13.0	366	2	F89936	conserved hypotet	230	6	13.0	573	1	HMS60	chaperonin groEL p
158	6	13.0	369	2	T22708	hypothetical prote	231	6	13.0	573	2	A34173	mitochondrial prot
159	6	13.0	370	2	H95173	hypothetical prote	232	6	13.0	575	2	E88651	protein B0212.2 [i
160	6	13.0	370	2	A90040	hypothetical prote	233	6	13.0	579	2	T47705	hypothetical prote
161	6	13.0	372	2	A61638	invasion protein i	234	6	13.0	583	2	T15644	probable cytochrom
162	6	13.0	372	2	AE0852	cell invasion prot	235	6	13.0	593	2	G87719	protein R119.7 [im
163	6	13.0	375	2	T09374	hypothetical prote	236	6	13.0	593	2	D97806	penicillin-binding
164	6	13.0	375	2	S73530	MG438 homolog E30	237	6	13.0	594	2	C71661	protein-tyrosine-p
165	6	13.0	379	1	JC2142	alkaline proteins	238	6	13.0	599	2	I51405	prolyl-tRNA synthe
166	6	13.0	381	2	G96804	hypothetical prote	239	6	13.0	604	2	AE2437	hypothetical prote
167	6	13.0	385	2	AF2427	hypothetical prote	240	6	13.0	612	2	T50226	hypothetical prote
168	6	13.0	387	2	A45117	aspartic proteins	241	6	13.0	612	2	T29761	hypothetical prote
169	6	13.0	393	2	T25590	hypothetical prote	242	6	13.0	617	2	S75447	proline-tRNA liga
170	6	13.0	393	2	AE0073	probable methytra	243	6	13.0	619	2	S29839	taurine transport
171	6	13.0	396	2	T47207	aspartic proteins	244	6	13.0	620	2	S46487	taurine transport
172	6	13.0	401	2	AE3537	oligopeptide trans	245	6	13.0	620	2	G01426	taurine transport
173	6	13.0	405	2	AE5832	probable glycosyl	246	6	13.0	634	2	E96927	probable permease
174	6	13.0	405	2	H64971	probable colanic a	247	6	13.0	638	2	S04640	methylmalonyl-CoA
175	6	13.0	405	2	F90986	probable glycosyl	248	6	13.0	639	2	E72336	maltodextrin glyco

249	6	13.0	639	2	S20887	actA protein precu	322	6	13.0	1787	2	G81684	excinuclease ABC c
250	6	13.0	639	2	A81100	actin-assembly ind	323	6	13.0	1807	2	JC6319	integrin beta-4 ch
251	6	13.0	649	2	H90163	DNA-directed RNA p	324	6	13.0	1852	2	A37860	calcium channel pr
252	6	13.0	662	2	S62707	NADH dehydrogenase	325	6	13.0	1855	2	A41649	DNA polymerase - m
253	6	13.0	664	2	A72215	excinuclease ABC c	326	6	13.0	1864	2	F86378	protein F21J9.12 [
254	6	13.0	671	2	T33431	hypothetical prote	327	6	13.0	1946	2	A82141	serine/threonine k
255	6	13.0	672	2	A11183	transcription regu	328	6	13.0	2022	2	T48818	glucan 1,4-alpha-g
256	6	13.0	688	2	AF1619	glycyl-tRNA synthe	329	6	13.0	2100	2	T38128	T7123.15 protein -
257	6	13.0	699	2	T15920	hypothetical prote	330	6	13.0	2139	2	A44467	voltage-dependent
258	6	13.0	702	2	E69498	hypothetical prote	331	6	13.0	2143	2	JH0427	voltage-dependent
259	6	13.0	715	2	T25631	hypothetical prote	332	6	13.0	2161	2	JH0564	calcium channel al
260	6	13.0	717	1	S28784	alpha-amylase (EC	333	6	13.0	2166	2	S11339	calcium channel pr
261	6	13.0	721	2	F70649	probable Acyl-CoA	334	6	13.0	2171	2	S05054	calcium channel al
262	6	13.0	726	2	A87334	ABC transporter, H	335	6	13.0	2181	2	A38198	calcium channel al
263	6	13.0	735	2	S10630	alpha subunit of c	336	6	13.0	2203	2	T42742	voltage-dependent
264	6	13.0	737	2	S72442	actin-fragmin kina	337	6	13.0	2220	2	A45290	calcium channel pr
265	6	13.0	758	2	AE2022	hypothetical prote	338	6	13.0	2271	2	F90073	hypothetical prote
266	6	13.0	759	2	S53922	PW6 protein - yea	339	6	13.0	2435	2	D69426	surface layer prot
267	6	13.0	804	2	T37821	probable dna repl	340	6	13.0	2693	2	A40743	IP3 receptor, XIP3
268	6	13.0	805	2	T21592	hypothetical prote	341	6	13.0	2747	2	B49132	fat facets (faf) s
269	6	13.0	817	2	T22442	hypothetical prote	342	6	13.0	2946	2	T15840	hypothetical prote
270	6	13.0	825	2	T27852	hypothetical prote	343	6	13.0	3194	2	D71917	toxin-like outer m
271	6	13.0	829	2	S12858	virA protein - Agr	344	6	13.0	3512	2	T17121	CPV protein - midg
272	6	13.0	830	2	A84951	DNA topoisomerase	345	6	13.0	4910	2	S64942	probable membrane
273	6	13.0	831	2	D96639	protein T1P9.12 [i	346	6	13.0	8243	2	T31307	type I fatty acid
274	6	13.0	833	2	S06972	virA protein - Agr	347	6	13.0	13288	2	T03099	mucin, submaxillar
275	6	13.0	833	2	AD3248	two component sens	348	6	13.0	26926	1	T38344	alkaline trypsin-I
276	6	13.0	841	2	A90669	probable enzyme [i	349	5	10.9	24	2	PN0653	hypothetical prote
277	6	13.0	841	2	C64755	yagX protein - Esc	350	5	10.9	26	2	S04376	myosin heavy chain
278	6	13.0	841	2	D85519	probable enzyme ya	351	5	10.9	36	2	B44400	hypothetical prote
279	6	13.0	842	2	T23715	hypothetical prote	352	5	10.9	36	2	H70243	hypothetical prote
280	6	13.0	857	2	A82768	protein-export mem	353	5	10.9	42	2	JE0005	actin alpha, smoot
281	6	13.0	874	2	G84644	probable ATP-depen	354	5	10.9	42	2	A99578	lipoprotein vsaH [
282	6	13.0	875	2	H90371	proteinase [import	355	5	10.9	44	2	C47193	1-caldesmon I (Hel
283	6	13.0	876	2	B37549	probable protein-e	356	5	10.9	45	1	HP80	haptoglobin precu
284	6	13.0	879	2	S73757	hypothetical prote	357	5	10.9	46	2	S42240	hypothetical prote
285	6	13.0	887	1	ROHYE	hydroxymethylgluta	358	5	10.9	46	2	I38225	protein-serine/thr
286	6	13.0	887	1	A23586	hydroxymethylgluta	359	5	10.9	46	4	A45758	hypothetical prote
287	6	13.0	888	1	RDHDE	hydroxymethylgluta	360	5	10.9	47	2	D81117	hypothetical prote
288	6	13.0	890	2	S41150	coat protein str	361	5	10.9	48	2	B64647	hypothetical prote
289	6	13.0	893	2	E72232	DNA-directed DNA p	362	5	10.9	50	2	F70227	hypothetical prote
290	6	13.0	968	2	T01733	hypothetical prote	363	5	10.9	51	2	A70229	hypothetical prote
291	6	13.0	970	2	G84939	hypothetical prote	364	5	10.9	51	2	G82455	hypothetical prote
292	6	13.0	971	2	T33907	hypothetical prote	365	5	10.9	52	2	D83682	hypothetical prote
293	6	13.0	977	2	B64593	hypothetical prote	366	5	10.9	53	2	S12520	core protein Al -
294	6	13.0	980	2	AD1922	metinyl-accepting c	367	5	10.9	53	2	F82662	hypothetical prote
295	6	13.0	1012	2	T43054	retinoblastoma-rel	368	5	10.9	53	2	AB1896	hypothetical prote
296	6	13.0	1032	2	T23164	hypothetical prote	369	5	10.9	54	2	I46468	actin - sheep (fra
297	6	13.0	1035	2	T23165	hypothetical prote	370	5	10.9	58	2	E97047	hypothetical prote
298	6	13.0	1049	2	T37933	hypothetical prote	371	5	10.9	60	1	LPSE5	antibiotic pep s p
299	6	13.0	1051	2	T51904	hypothetical prote	372	5	10.9	62	2	A65045	hypothetical prote
300	6	13.0	1052	2	A22959	conserved hypothet	373	5	10.9	62	2	E63346	hypothetical prote
301	6	13.0	1054	2	D96519	myosin-like protein	374	5	10.9	63	2	T12825	hypothetical prote
302	6	13.0	1072	2	G95851	probable hemolysin	375	5	10.9	64	2	A86333	hypothetical prote
303	6	13.0	1075	2	T38253	hypothetical prote	376	5	10.9	65	2	S73176	ribosomal protein
304	6	13.0	1110	2	I51116	NF-180 - sea lamp	377	5	10.9	66	2	T29601	hypothetical prote
305	6	13.0	1126	2	S04716	DNA-directed RNA p	378	5	10.9	67	2	I46473	alpha-actin - rabb
306	6	13.0	1160	2	T23713	hypothetical prote	379	5	10.9	67	2	C86830	hypothetical prote
307	6	13.0	1164	2	G71827	hypothetical prote	380	5	10.9	67	2	E81135	hypothetical prote
308	6	13.0	1184	2	A56638	hypothetical prote	381	5	10.9	67	2	E81042	hypothetical prote
309	6	13.0	1286	2	T23714	hypothetical prote	382	5	10.9	67	2	E69447	hypothetical prote
310	6	13.0	1291	2	T17242	hypothetical prote	383	5	10.9	67	2	T07256	hypothetical prote
311	6	13.0	1331	2	S05011	calcium channel al	384	5	10.9	69	1	MIEC77	microcin B17 precu
312	6	13.0	1341	2	H98333	hypothetical prote	385	5	10.9	72	2	A35511	homeotic protein H
313	6	13.0	1355	2	A47026	alkaline phosphata	386	5	10.9	72	2	S75519	hypothetical prote
314	6	13.0	1373	2	E97722	hypothetical prote	387	5	10.9	73	1	HP88	haptoglobin precu
315	6	13.0	1374	2	A71724	dna-directed RNA p	388	5	10.9	75	2	H81320	small hydrophobic
316	6	13.0	1529	2	T02750	RNA directed DNA p	389	5	10.9	76	2	B75271	hypothetical prote
317	6	13.0	1559	2	T30535	calcium channel al	390	5	10.9	77	2	T14995	hypothetical prote
318	6	13.0	1610	2	A46227	voltage-dependent	391	5	10.9	78	2	T31106	probable glycerol-
319	6	13.0	1642	2	T08880	NMDA receptor-bind	392	5	10.9	78	2	C71978	hypothetical prote
320	6	13.0	1646	2	JH0422	voltage-dependent	393	5	10.9	78	2	SI7872	translation initia
321	6	13.0	1748	1	JN0786	integrin beta-4 ch	394	5	10.9	79	2	H82705	hypothetical prote

395	5	10.9	79	2	AH3369	hypothetical prote	468	5	10.9	102	2	G83721	hypothetical prote
396	5	10.9	80	2	E71430	hypothetical prote	469	5	10.9	102	2	AD1748	B ₁ subtilis yfht p
397	5	10.9	81	2	C97140	hypothetical prote	470	5	10.9	103	2	A46011	cysteine proteinas
398	5	10.9	82	2	T18102	hypothetical prote	471	5	10.9	103	2	A86584	Cro311 hypothetical
399	5	10.9	83	2	A26353	conserved hypothet	472	5	10.9	104	2	S53942	probable membrane
400	5	10.9	84	2	A46348	hypothetical prote	473	5	10.9	104	2	B71450	hypothetical prote
401	5	10.9	85	2	A96949	hypothetical prote	474	5	10.9	104	2	S51479	drought-induced pr
402	5	10.9	86	2	F98226	hypothetical prote	475	5	10.9	104	2	F86867	prophage ps3 prote
403	5	10.9	87	2	AB3060	conserved hypothet	476	5	10.9	104	2	AG3518	hypothetical prote
404	5	10.9	88	2	AC3378	hypothetical membr	477	5	10.9	105	2	T17976	hypothetical prote
405	5	10.9	89	2	S53038	hypothetical prote	478	5	10.9	105	2	A82914	hypothetical prote
406	5	10.9	90	2	E90475	hypothetical prote	479	5	10.9	105	2	H32537	T-cell receptor be
407	5	10.9	91	2	F84014	hypothetical prote	480	5	10.9	106	2	A03859	hypothetical prote
408	5	10.9	92	2	AF2777	conserved hypothet	481	5	10.9	106	2	H69310	hypothetical prote
409	5	10.9	93	2	T26337	hypothetical prote	482	5	10.9	106	2	D70212	hypothetical prote
410	5	10.9	94	2	A64337	flavocytochrome c,	483	5	10.9	106	2	G97841	acetyl-CoA acetyl
411	5	10.9	95	2	D71386	probable ribosomal	484	5	10.9	107	2	D71114	hypothetical prote
412	5	10.9	96	2	B97266	hypothetical prote	485	5	10.9	107	2	JC1127	major allergen cha
413	5	10.9	97	2	AC0730	hypothetical prote	486	5	10.9	107	2	B85356	glycine-rich prote
414	5	10.9	98	2	XV8BD	Shigella toxin cha	487	5	10.9	107	2	E69318	hypothetical prote
415	5	10.9	99	2	JN0726	Shiga-like toxin I	488	5	10.9	108	2	T45320	hypothetical prote
416	5	10.9	100	2	I56163	calgranulin A - mo	489	5	10.9	108	2	D70711	hypothetical prote
417	5	10.9	101	2	XV8PH9	Shiga-like toxin I	490	5	10.9	108	2	F64010	hypothetical prote
418	5	10.9	102	2	B53887	Shiga toxin I subu	491	5	10.9	109	2	AB0218	probable acyl carr
419	5	10.9	103	2	E91000	hypothetical prote	492	5	10.9	109	2	S31978	Ig kappa chain - h
420	5	10.9	104	2	K3141	hypothetical prote	493	5	10.9	109	2	C56413	hypothetical prote
421	5	10.9	105	2	G85845	Shiga toxin I subu	494	5	10.9	109	2	T35249	quaternary ammoniu
422	5	10.9	106	2	A54937	brain type III sod	495	5	10.9	110	2	AG3382	csqC protein precu
423	5	10.9	107	2	P90555	hypothetical prote	496	5	10.9	110	2	S70789	probable curli pro
424	5	10.9	108	2	I54248	apolipoprotein A-I	497	5	10.9	110	2	E90806	probable curli pro
425	5	10.9	109	2	G46028	hypothetical prote	498	5	10.9	110	2	A85666	YbL protein precu
426	5	10.9	110	2	B45837	Ig gamma-1 chain C	499	5	10.9	110	2	B64889	hypothetical 13.1k
427	5	10.9	111	2	I55688	seminal vesicle pr	500	5	10.9	111	2	S56499	hypothetical prote
428	5	10.9	112	2	S57260	nitrite reductase	501	5	10.9	112	2	T31580	hypothetical prote
429	5	10.9	113	2	A81105	probable periplasm	502	5	10.9	112	2	T26105	hypothetical prote
430	5	10.9	114	2	D82451	conserved hypothet	503	5	10.9	113	2	J01128	gas-vesicle operon
431	5	10.9	115	2	AE2080	hypothetical prote	504	5	10.9	113	2	T08234	gas-vesicle operon
432	5	10.9	116	2	AE3173	transcription regu	505	5	10.9	113	2	A83697	hypothetical prote
433	5	10.9	117	2	C63684	ribosomal protein	506	5	10.9	113	2	I51025	transcription fact
434	5	10.9	118	2	C71835	ribosomal protein	507	5	10.9	113	2	F81370	hypothetical prote
435	5	10.9	119	2	A10757	probable cobalt tr	508	5	10.9	113	2	JN0784	acid phosphatase,
436	5	10.9	120	2	AG1359	holin (Bacterioph	509	5	10.9	114	2	F71141	hypothetical prote
437	5	10.9	121	2	AH1448	holin (Bacterioph	510	5	10.9	115	2	T30042	hypothetical prote
438	5	10.9	122	2	H42518	Al5L 11k protein -	511	5	10.9	115	2	T12916	hypothetical prote
439	5	10.9	123	2	G36849	Al5L protein - var	512	5	10.9	115	2	E86809	hypothetical prote
440	5	10.9	124	2	T28557	hypothetical prote	513	5	10.9	116	2	H69023	conserved hypothet
441	5	10.9	125	2	F72165	Al6L protein - var	514	5	10.9	116	2	AI2446	hypothetical prote
442	5	10.9	126	2	S70353	KIN28 protein homo	515	5	10.9	116	2	C72232	hypothetical prote
443	5	10.9	127	2	A70803	early secretory an	516	5	10.9	117	1	A59316	ghrelin precursor
444	5	10.9	128	2	C69449	conserved hypothet	517	5	10.9	117	2	F72620	hypothetical prote
445	5	10.9	129	2	S69800	holin protein 10.3	518	5	10.9	117	2	S28124	gas-vesicle operon
446	5	10.9	130	2	S69798	holin protein, 10.	519	5	10.9	117	2	JX0053	lysozyme (EC 3.2.1
447	5	10.9	131	2	S26652	Ig gamma-1 chain C	520	5	10.9	117	2	T49511	hypothetical prote
448	5	10.9	132	2	E95064	conserved hypothet	521	5	10.9	118	2	JC2568	hypothetical prote
449	5	10.9	133	2	C90661	hypothetical prote	522	5	10.9	118	2	T06659	histone H2A.T6G15
450	5	10.9	134	2	A64748	yafN protein - Esc	523	5	10.9	118	2	T28246	hypothetical prote
451	5	10.9	135	2	G64454	hypothetical prote	524	5	10.9	118	2	T30509	gas vesicle protei
452	5	10.9	136	2	F84155	blb operon transc	525	5	10.9	119	2	S28133	conserved hypothet
453	5	10.9	137	2	A85312	hypothetical prote	526	5	10.9	119	2	B90168	hypothetical prote
454	5	10.9	138	2	G97931	conserved hypothet	527	5	10.9	119	2	H70437	ribosome-binding f
455	5	10.9	139	2	AH0965	probable secreted	528	5	10.9	120	2	H84973	hypothetical prote
456	5	10.9	140	2	T42307	hypothetical prote	529	5	10.9	120	2	F71104	hypothetical prote
457	5	10.9	141	2	AH1666	integral membrane	530	5	10.9	120	2	B71020	hypothetical prote
458	5	10.9	142	2	AC0152	hypothetical prote	531	5	10.9	120	2	G71075	hypothetical prote
459	5	10.9	143	2	AB3452	urease operon 23k	532	5	10.9	122	1	A42628	Ig V-P-J region (M
460	5	10.9	144	2	SL4851	actin - fruit fly	533	5	10.9	122	2	S69909	probable ribosomal
461	5	10.9	145	2	A87463	hypothetical prote	534	5	10.9	122	2	T41471	seminal vesicle se
462	5	10.9	146	2	T45066	H+-transporting AT	535	5	10.9	122	2	S29540	insertion element
463	5	10.9	147	2	I38611	calcium channel, v	536	5	10.9	122	2	G81855	hypothetical prote
464	5	10.9	148	2	H81352	small hydrophobic	537	5	10.9	122	2	AB1199	hypothetical prote
465	5	10.9	149	2	S51974	probable membrane	538	5	10.9	122	2	AC1848	hypothetical prote
466	5	10.9	150	2	A97534	hypothetical prote	539	5	10.9	123	2	D72519	hypothetical prote
467	5	10.9	151	2			540	5	10.9	123	2		

541	5	10.9	123	2	G72581	hypothetical prote	614	5	10.9	146	2	T06471	core protein - gar
542	5	10.9	123	2	F83207	hypothetical prote	615	5	10.9	146	2	C98889	conserved hypothet
543	5	10.9	124	2	S03521	Ig kappa chain pre	616	5	10.9	147	1	A59315	proteinase synthas
544	5	10.9	125	2	A26337	actin, skeletal mu	617	5	10.9	147	2	D90126	ubiquitin-conjugat
545	5	10.9	125	2	C83138	hypothetical prote	618	5	10.9	147	2	AH1196	B. subtilis Yjcf p
546	5	10.9	125	2	T38515	hypothetical prote	619	5	10.9	147	2	A81454	probable integral
547	5	10.9	125	2	AG1653	Antigen C homolog	620	5	10.9	147	2	A85554	conserved hypothet
548	5	10.9	126	2	S50574	hypothetical prote	621	5	10.9	147	2	B69040	hypothetical prote
549	5	10.9	128	2	AB2487	hypothetical prote	622	5	10.9	147	2	C70312	hypothetical prote
550	5	10.9	128	2	T32789	hypothetical prote	623	5	10.9	147	2	F83801	chorismate mutase
551	5	10.9	128	2	AE1701	hypothetical prote	624	5	10.9	147	2	A39145	hypothetical prote
552	5	10.9	129	2	F72488	hypothetical prote	625	5	10.9	147	2	E64370	hypothetical prote
553	5	10.9	129	2	E90535	hypothetical prote	626	5	10.9	147	2	A39091	phospholipase A2 i
554	5	10.9	129	2	B90175	NADH dehydrogenase	627	5	10.9	147	2	B39091	phospholipase A2 i
555	5	10.9	130	2	PL0098	Ig heavy chain pre	628	5	10.9	147	2	A75093	hypothetical prote
556	5	10.9	130	2	G71212	hypothetical prote	629	5	10.9	147	2	F71050	hypothetical prote
557	5	10.9	131	2	S67869	probable membrane	630	5	10.9	148	2	E64939	hypothetical prote
558	5	10.9	131	2	S63305	probable membrane	631	5	10.9	148	2	B90941	hypothetical prote
559	5	10.9	132	2	B69489	LSU ribosomal prot	632	5	10.9	148	2	F85789	hypothetical prote
560	5	10.9	132	2	AC3143	hypothetical prote	633	5	10.9	148	2	D87687	conserved hypothet
561	5	10.9	133	2	D69836	hypothetical prote	634	5	10.9	148	2	S10655	hypothetical prote
562	5	10.9	133	2	F71189	hypothetical prote	635	5	10.9	148	2	A36641	hypothetical prote
563	5	10.9	133	2	C97557	hypothetical prote	636	5	10.9	148	2	T52550	probable transcrip
564	5	10.9	134	2	F75170	hypothetical prote	637	5	10.9	148	2	AG1122	hypothetical prote
565	5	10.9	134	2	T43561	probable transposa	638	5	10.9	148	2	AI2051	two-component resp
566	5	10.9	134	2	S72727	deda protein - Myc	639	5	10.9	149	2	C95188	PTS system, IIA co
567	5	10.9	135	2	S38486	T-cell receptor be	640	5	10.9	149	2	H72344	conserved hypothet
568	5	10.9	135	2	T49327	cofilin related pr	641	5	10.9	149	2	D84693	hypothetical prote
569	5	10.9	135	2	G64600	hypothetical prote	642	5	10.9	149	2	C90824	hypothetical membr
570	5	10.9	135	2	A36828	orf10 protein - Au	643	5	10.9	150	1	MOCHA2	myosin alkali ligh
571	5	10.9	135	2	G86695	hypothetical prote	644	5	10.9	150	2	G85939	type III secretion
572	5	10.9	135	2	T31350	hypothetical prote	645	5	10.9	151	1	MGUR2C	Spec2c protein - s
573	5	10.9	136	2	T02870	globulin 2 precurs	646	5	10.9	151	2	T42298	hypothetical prote
574	5	10.9	136	2	AG2009	transcription regu	647	5	10.9	151	2	AD1569	hypothetical prote
575	5	10.9	136	2	T37767	probable transcrip	648	5	10.9	151	2	AG1215	hypothetical prote
576	5	10.9	137	2	A28258	actin 5C - fruit f	649	5	10.9	151	2	D87548	hypothetical prote
577	5	10.9	137	2	T04930	glycine-rich cell	650	5	10.9	151	2	B86474	hypothetical prote
578	5	10.9	137	2	AB2544	transcription regu	651	5	10.9	152	2	B98851	hypothetical prote
579	5	10.9	137	2	G72666	hypothetical prote	652	5	10.9	152	2	S72685	hypothetical prote
580	5	10.9	138	1	A69219	conserved hypothet	653	5	10.9	153	2	A82431	hypothetical prote
581	5	10.9	138	2	S67968	transition protein	654	5	10.9	153	2	F90704	hypothetical prote
582	5	10.9	138	2	H72452	hypothetical prote	655	5	10.9	153	2	H72777	hypothetical prote
583	5	10.9	138	2	AE1324	PTS mannose-specif	656	5	10.9	153	2	H83895	transcription regu
584	5	10.9	139	2	A71190	hypothetical prote	657	5	10.9	153	2	A85555	hypothetical prote
585	5	10.9	140	2	S73680	MG337 homolog p02	658	5	10.9	154	2	JS0710	heat shock protein
586	5	10.9	140	2	H69405	hypothetical prote	659	5	10.9	154	2	S24396	heat shock protein
587	5	10.9	140	2	J02293	transcription enha	660	5	10.9	154	2	T44801	heat shock protein
588	5	10.9	141	2	A84457	40S ribosomal prot	661	5	10.9	154	2	H82810	bacterioferritin X
589	5	10.9	141	2	T48476	40S ribosomal prot	662	5	10.9	154	2	H75031	hypothetical prote
590	5	10.9	141	2	S30832	hypothetical prote	663	5	10.9	154	2	D83516	hypothetical prote
591	5	10.9	141	2	F71176	hypothetical prote	664	5	10.9	154	2	G89932	hypothetical prote
592	5	10.9	141	2	B64399	hypothetical prote	665	5	10.9	155	2	A60945	actin homolog FAT,
593	5	10.9	142	2	T08343	arsC protein - Hal	666	5	10.9	155	2	F71004	hypothetical prote
594	5	10.9	142	2	T46832	heat shock protein	667	5	10.9	155	2	A71693	hypothetical prote
595	5	10.9	142	2	F97079	PTS system, fructo	668	5	10.9	155	2	C91094	type III secretion
596	5	10.9	143	2	S17939	xpsG protein - Xan	669	5	10.9	156	1	A49342	acetyl-CoA carboxy
597	5	10.9	143	2	D72785	hcv homolog AP802	670	5	10.9	156	1	D69019	conserved hypothet
598	5	10.9	143	2	T28888	hypothetical prote	671	5	10.9	156	2	F86661	acetyltransferase
599	5	10.9	143	2	G72000	hth transcription	672	5	10.9	156	2	T16531	hypothetical prote
600	5	10.9	143	2	B86624	HTH transcription	673	5	10.9	157	2	A71137	hypothetical prote
601	5	10.9	143	2	T49689	hypothetical prote	674	5	10.9	157	2	S75191	hypothetical prote
602	5	10.9	143	2	JC7309	agglutinin I - Jer	675	5	10.9	157	2	G97760	hypothetical prote
603	5	10.9	144	2	T33494	hypothetical prote	676	5	10.9	157	2	AG0537	probable secreted
604	5	10.9	144	2	H71853	hypothetical prote	677	5	10.9	157	2	T46178	hypothetical prote
605	5	10.9	144	2	H64659	hypothetical prote	678	5	10.9	157	2	T46440	hypothetical prote
606	5	10.9	144	2	C81910	probable periplasm	679	5	10.9	158	2	D82311	conserved hypothet
607	5	10.9	145	2	AF1554	B. subtilis Yjcf p	680	5	10.9	158	2	B82191	hypothetical prote
608	5	10.9	145	2	A47888	hemagglutinating s	681	5	10.9	158	2	F83475	conserved hypothet
609	5	10.9	145	2	T15089	hypothetical prote	682	5	10.9	158	2	T25654	hypothetical prote
610	5	10.9	145	2	A87526	hypothetical prote	683	5	10.9	159	2	A43599	cytolysin IIC - Ac
611	5	10.9	145	2	T39564	hypothetical prote	684	5	10.9	159	2	S15640	phlIn (variant Ngr
612	5	10.9	145	2	C82513	conserved hypothet	685	5	10.9	159	2	G82924	hypothetical u0180
613	5	10.9	145	2	T36091	probable maturase-	686	5	10.9	159	2	D95261	conserved hypothet

687 conserved hypochet 159 2 A98127
688 2C-methyl-D-erythr 159 2 I55083
689 2C-methyl-D-erythr 159 2 H91078
690 2C-methyl-D-erythr 159 2 A85924
691 2C-methyl-D-erythr 159 2 A08056
692 2C-methyl-D-erythr 159 2 S15964
693 2C-methyl-D-erythr 159 2 E95328
694 2C-methyl-D-erythr 159 2 G88082
695 2C-methyl-D-erythr 159 2 G84125
696 2C-methyl-D-erythr 159 2 H43834
697 2C-methyl-D-erythr 159 2 H82207
698 2C-methyl-D-erythr 159 2 T32026
699 2C-methyl-D-erythr 159 2 S26039
700 2C-methyl-D-erythr 159 2 B71687
701 2C-methyl-D-erythr 159 2 G64355
702 2C-methyl-D-erythr 159 2 F70006
703 2C-methyl-D-erythr 159 2 B90417
704 2C-methyl-D-erythr 159 2 AD0836
705 2C-methyl-D-erythr 159 2 AD0930
706 2C-methyl-D-erythr 159 2 B86421
707 2C-methyl-D-erythr 159 2 AE1668
708 2C-methyl-D-erythr 159 2 AG1296
709 2C-methyl-D-erythr 159 2 C85356
710 2C-methyl-D-erythr 159 2 T35903
711 2C-methyl-D-erythr 159 2 F86637
712 2C-methyl-D-erythr 159 2 S65764
713 2C-methyl-D-erythr 159 2 AC3839
714 2C-methyl-D-erythr 159 2 A73160
715 2C-methyl-D-erythr 159 2 A70369
716 2C-methyl-D-erythr 159 2 G81183
717 2C-methyl-D-erythr 159 2 I40526
718 2C-methyl-D-erythr 159 2 I40183
719 2C-methyl-D-erythr 159 2 AC0537
720 2C-methyl-D-erythr 159 2 T01777
721 2C-methyl-D-erythr 159 2 PS0261
722 2C-methyl-D-erythr 159 2 B82670
723 2C-methyl-D-erythr 159 2 AH1249
724 2C-methyl-D-erythr 159 2 KR022B
725 2C-methyl-D-erythr 159 2 AE2392
726 2C-methyl-D-erythr 159 2 AG3566
727 2C-methyl-D-erythr 159 2 D72394
728 2C-methyl-D-erythr 159 2 F97128
729 2C-methyl-D-erythr 159 2 A11444
730 2C-methyl-D-erythr 159 2 B81676
731 2C-methyl-D-erythr 159 2 B71553
732 2C-methyl-D-erythr 159 2 F84559
733 2C-methyl-D-erythr 159 2 E64367
734 2C-methyl-D-erythr 159 2 T25892
735 2C-methyl-D-erythr 159 2 G82898
736 2C-methyl-D-erythr 159 2 S4745
737 2C-methyl-D-erythr 159 2 F82403
738 2C-methyl-D-erythr 159 2 C84923
739 2C-methyl-D-erythr 159 2 B87602
740 2C-methyl-D-erythr 159 2 D86801
741 2C-methyl-D-erythr 159 2 S33391
742 2C-methyl-D-erythr 159 2 S34127
743 2C-methyl-D-erythr 159 2 JC1220
744 2C-methyl-D-erythr 159 2 A35944
745 2C-methyl-D-erythr 159 2 JC7242
746 2C-methyl-D-erythr 159 2 JC2213
747 2C-methyl-D-erythr 159 2 T27263
748 2C-methyl-D-erythr 159 2 F82908
749 2C-methyl-D-erythr 159 2 C90079
750 2C-methyl-D-erythr 159 2 T28000
751 2C-methyl-D-erythr 159 2 T07380
752 2C-methyl-D-erythr 159 2 E69711
753 2C-methyl-D-erythr 159 2 H84709
754 2C-methyl-D-erythr 159 2 S66014
755 2C-methyl-D-erythr 159 2 H83542
756 2C-methyl-D-erythr 159 2 F83585
757 2C-methyl-D-erythr 159 2 A70864
758 2C-methyl-D-erythr 159 2 S18035
759 2C-methyl-D-erythr 159 2 S18034

760 conserved hypochet 10.9 5
761 2C-methyl-D-erythr 10.9 5
762 2C-methyl-D-erythr 10.9 5
763 2C-methyl-D-erythr 10.9 5
764 2C-methyl-D-erythr 10.9 5
765 2C-methyl-D-erythr 10.9 5
766 2C-methyl-D-erythr 10.9 5
767 2C-methyl-D-erythr 10.9 5
768 2C-methyl-D-erythr 10.9 5
769 2C-methyl-D-erythr 10.9 5
770 2C-methyl-D-erythr 10.9 5
771 2C-methyl-D-erythr 10.9 5
772 2C-methyl-D-erythr 10.9 5
773 2C-methyl-D-erythr 10.9 5
774 2C-methyl-D-erythr 10.9 5
775 2C-methyl-D-erythr 10.9 5
776 2C-methyl-D-erythr 10.9 5
777 2C-methyl-D-erythr 10.9 5
778 2C-methyl-D-erythr 10.9 5
779 2C-methyl-D-erythr 10.9 5
780 2C-methyl-D-erythr 10.9 5
781 2C-methyl-D-erythr 10.9 5
782 2C-methyl-D-erythr 10.9 5
783 2C-methyl-D-erythr 10.9 5
784 2C-methyl-D-erythr 10.9 5
785 2C-methyl-D-erythr 10.9 5
786 2C-methyl-D-erythr 10.9 5
787 2C-methyl-D-erythr 10.9 5
788 2C-methyl-D-erythr 10.9 5
789 2C-methyl-D-erythr 10.9 5
790 2C-methyl-D-erythr 10.9 5
791 2C-methyl-D-erythr 10.9 5
792 2C-methyl-D-erythr 10.9 5
793 2C-methyl-D-erythr 10.9 5
794 2C-methyl-D-erythr 10.9 5
795 2C-methyl-D-erythr 10.9 5
796 2C-methyl-D-erythr 10.9 5
797 2C-methyl-D-erythr 10.9 5
798 2C-methyl-D-erythr 10.9 5
799 2C-methyl-D-erythr 10.9 5
800 2C-methyl-D-erythr 10.9 5
801 2C-methyl-D-erythr 10.9 5
802 2C-methyl-D-erythr 10.9 5
803 2C-methyl-D-erythr 10.9 5
804 2C-methyl-D-erythr 10.9 5
805 2C-methyl-D-erythr 10.9 5
806 2C-methyl-D-erythr 10.9 5
807 2C-methyl-D-erythr 10.9 5
808 2C-methyl-D-erythr 10.9 5
809 2C-methyl-D-erythr 10.9 5
810 2C-methyl-D-erythr 10.9 5
811 2C-methyl-D-erythr 10.9 5
812 2C-methyl-D-erythr 10.9 5
813 2C-methyl-D-erythr 10.9 5
814 2C-methyl-D-erythr 10.9 5
815 2C-methyl-D-erythr 10.9 5
816 2C-methyl-D-erythr 10.9 5
817 2C-methyl-D-erythr 10.9 5
818 2C-methyl-D-erythr 10.9 5
819 2C-methyl-D-erythr 10.9 5
820 2C-methyl-D-erythr 10.9 5
821 2C-methyl-D-erythr 10.9 5
822 2C-methyl-D-erythr 10.9 5
823 2C-methyl-D-erythr 10.9 5
824 2C-methyl-D-erythr 10.9 5
825 2C-methyl-D-erythr 10.9 5
826 2C-methyl-D-erythr 10.9 5
827 2C-methyl-D-erythr 10.9 5
828 2C-methyl-D-erythr 10.9 5
829 2C-methyl-D-erythr 10.9 5
830 2C-methyl-D-erythr 10.9 5
831 2C-methyl-D-erythr 10.9 5
832 2C-methyl-D-erythr 10.9 5

thamatin-like pro
hypothetical prote
hypothetical prote
hypothetical prote
PT2-F1 steroid rec
probable transcrip
hypothetical prote
myelin basic prote
inositol 1,4,5-tri
heat shock protein
hypothetical prote
conserved hypotet
conserved hypotet
hypothetical prote
MHC class I protei
calcineurin regula
hypothetical prote
ribosomal protein
fimbrial protein
conserved hypotet
hypothetical prote
conserved hypotet
conserved hypotet
thymidine kinase
parb protein homol
ail protein precur
hypothetical prote
ribosomal protein
secretory granule
calcineurin regula
conserved hypotet
40S ribosomal prot
surface antigen fu
hypothetical prote
hypothetical prote
hypothetical prote
probable adenine p
G-protein signalin
Syd protein fmpor
hypothetical prote
K99 fimbrial subun
keratin, 68K type
senescence-associa
hypothetical prote
glycine-rich cell
senescence-associa
lema protein (U661
conserved hypotet
hypothetical prote
probable phase-rel
adenylate kinase
probable membrane
hypothetical prote
hypothetical prote
ATP synthase Fl, d
hypothetical prote
probable dirigent
dihydrofolate redu
dihydrofolate redu
hypothetical prote
dihydrofolate redu
siroheme synthase
aspartic proteinas
histidinol-phospha
probable phosphata
hypothetical prote
hypothetical prote
conserved hypotet
probable phosphono
hypothetical prote

833	5	10.9	188	2	A11225	hypothetical prote	906	5	10.9	202	2	AG1643	1-acylglycerol-3-p
834	5	10.9	188	2	B87875	conserved hypotet	907	5	10.9	203	1	WBFBTB	gene 10 protein
835	5	10.9	189	2	A86699	thymidine kinase (908	5	10.9	203	2	D81328	probable riboflavi
836	5	10.9	189	2	JC4072	virulence-associat	909	5	10.9	203	2	D81700	thymidylate kinase
837	5	10.9	189	2	F82448	conserved hypotet	910	5	10.9	203	2	JC7521	ribosomal protein
838	5	10.9	189	2	AD1224	precorrin decarbox	911	5	10.9	203	2	T02137	hypothetical prote
839	5	10.9	189	2	AF1577	conserved hypotet	912	5	10.9	203	2	E64332	signal peptidase I
840	5	10.9	189	2	B95219	conserved hypotet	913	5	10.9	203	2	F90886	H repeat-associate
841	5	10.9	189	2	H98082	conserved hypotet	914	5	10.9	203	2	F90302	H repeat-associate
842	5	10.9	189	2	D62123	hypothetical prote	915	5	10.9	203	2	C85731	actin (clone Tac9)
843	5	10.9	189	2	D69128	conserved hypotet	916	5	10.9	204	2	S35256	hypothetical prote
844	5	10.9	189	2	E97498	outer membrane ser	917	5	10.9	204	2	T05677	recF protein limpo
845	5	10.9	190	2	T06953	anthranilate synth	918	5	10.9	205	2	T46553	hypothetical prote
846	5	10.9	190	2	D75401	transcription repr	919	5	10.9	205	2	A84717	transcription regu
847	5	10.9	190	2	H59105	hypothetical prote	920	5	10.9	206	2	B97302	hypothetical prote
848	5	10.9	190	2	T26019	hypothetical prote	921	5	10.9	206	2	A40305	biliary glycoprote
849	5	10.9	191	2	AB1768	thymidine kinase h	922	5	10.9	206	2	A81231	50S ribosomal prot
850	5	10.9	191	2	AH1392	thymidine kinase h	923	5	10.9	206	2	D68832	uridine kinase (PC
851	5	10.9	191	2	T19804	hypothetical prote	924	5	10.9	206	2	T38724	probable peroxisom
852	5	10.9	191	2	T19803	hypothetical prote	925	5	10.9	206	2	AG2069	hypothetical prote
853	5	10.9	192	1	A64627	phosphoheptose iso	926	5	10.9	206	2	G86965	conserved hypotet
854	5	10.9	192	2	B71887	phosphoheptose iso	927	5	10.9	206	2	F64620	hypothetical prote
855	5	10.9	192	2	S70779	asparagine--trNA I	928	5	10.9	206	2	E71894	hypothetical prote
856	5	10.9	192	2	G71089	hypothetical prote	929	5	10.9	206	2	B28439	maturase-related h
857	5	10.9	193	1	B69043	conserved hypotet	930	5	10.9	207	2	T07381	glycine-rich prote
858	5	10.9	193	2	D72474	probable deoxy-ribo	931	5	10.9	207	2	S67776	probable membrane
859	5	10.9	193	2	F64158	hypothetical prote	932	5	10.9	207	2	G70101	uridine kinase (ud
860	5	10.9	193	2	AF1981	hypothetical prote	933	5	10.9	207	2	B89943	uridine kinase (im
861	5	10.9	193	2	T28200	hypothetical prote	934	5	10.9	207	2	D83187	hypothetical prote
862	5	10.9	193	2	AG2389	hypothetical prote	935	5	10.9	207	2	AH1190	3-methyladenine DN
863	5	10.9	193	2	S18362	probable membrane	936	5	10.9	207	2	AH1548	3-methyladenine DN
864	5	10.9	194	2	T16336	icmc protein - Leg	937	5	10.9	207	2	C69780	hypothetical prote
865	5	10.9	194	2	AD1754	proteinase (bacter	938	5	10.9	207	2	G84983	hypothetical prote
866	5	10.9	194	2	AI0241	hypothetical prote	939	5	10.9	208	2	T46896	5-amino-6-(5-phosp
867	5	10.9	195	2	S39777	actin beta - pig (940	5	10.9	208	2	D88961	merozoite surface
868	5	10.9	195	2	T21997	hypothetical prote	941	5	10.9	208	2	B86874	protein F59A7.6 (i
869	5	10.9	195	2	AD2782	hypothetical prote	942	5	10.9	208	2	B86874	hypothetical prote
870	5	10.9	195	2	A99346	hypothetical prote	943	5	10.9	208	2	AB1485	hypothetical prote
871	5	10.9	195	2	D70571	hypothetical prote	944	5	10.9	209	2	AC1624	hypothetical prote
872	5	10.9	196	1	AQ3RNV	aequorin precursor	945	5	10.9	209	2	AI1261	uridine kinase hom
873	5	10.9	196	2	A30327	corticolliberin pre	946	5	10.9	209	2	AI1261	uridine kinase hom
874	5	10.9	196	2	E97561	hypothetical prote	947	5	10.9	210	2	H82037	gp51 (bacteriophag
875	5	10.9	196	2	E90273	conserved hypotet	948	5	10.9	210	2	S49577	conserved hypotet
876	5	10.9	196	2	G65039	hypothetical prote	949	5	10.9	210	2	E75553	chloramphenicol O-
877	5	10.9	196	2	B71232	hypothetical prote	950	5	10.9	210	2	E84121	uridine kinase - D
878	5	10.9	197	2	F71248	probable proteasom	951	5	10.9	210	2	F69250	stage II sporulati
879	5	10.9	197	2	H75212	probable proteasom	952	5	10.9	210	2	T40465	corrinoid methyltr
880	5	10.9	197	2	T03442	glycine-rich prote	953	5	10.9	211	2	G59728	probable mitosis a
881	5	10.9	197	2	T43558	hypothetical prote	954	5	10.9	211	2	T36373	uridine kinase udk
882	5	10.9	197	2	C64422	hypothetical prote	955	5	10.9	211	2	S08522	hypothetical prote
883	5	10.9	197	2	D97061	hypothetical prote	956	5	10.9	211	2	T19859	probable sodium-de
884	5	10.9	197	2	AH1372	uncharacterized lo	957	5	10.9	211	2	T41218	hypothetical prote
885	5	10.9	198	2	A57716	proteins involved	958	5	10.9	211	2	AE1711	hypothetical prote
886	5	10.9	198	2	T06813	thiol-specific ant	959	5	10.9	212	2	AG0327	hypothetical prote
887	5	10.9	198	2	G97088	dehydrin 3 - garde	960	5	10.9	212	2	B81669	conserved hypotet
888	5	10.9	198	2	AG3239	amidase, related t	961	5	10.9	212	2	B82920	hypothetical prote
889	5	10.9	199	2	B97987	cell filamentation	962	5	10.9	213	1	RB8EHS	hypothetical prote
890	5	10.9	199	2	D95117	thymidine kinase (963	5	10.9	213	2	AG9310	proteasome, subuni
891	5	10.9	199	2	D36131	sigma-B activity i	964	5	10.9	213	2	A21177	Ig light chain pre
892	5	10.9	199	2	T25260	hypothetical prote	965	5	10.9	213	2	S68213	Ig heavy chain (Ma
893	5	10.9	199	2	B86655	conserved hypotet	966	5	10.9	213	2	AG1043	Ig heavy chain (Ma
894	5	10.9	199	2	AD1271	probable periplasm	967	5	10.9	213	2	S63996	actin CAL5 - sea s
895	5	10.9	199	2	AD1230	phosphoglycerate m	968	5	10.9	213	2	S06788	pilin precursor -
896	5	10.9	199	2	AG1583	weakly phosphoglyc	969	5	10.9	214	2	S17680	fibrillar protein h
897	5	10.9	200	2	E85047	hypothetical prote	970	5	10.9	214	2	PC4202	fibrinolytic prote
898	5	10.9	200	2	B82483	hypothetical prote	971	5	10.9	214	2	T10764	monoclonal antibod
899	5	10.9	200	2	B69168	hypothetical prote	972	5	10.9	214	2	C86967	probable MADS box
900	5	10.9	200	2	D95950	probable esterase	973	5	10.9	214	2	B96755	probable membrane
901	5	10.9	201	2	A31484	troponin I, fast s	974	5	10.9	214	2	F86197	hypothetical prote
902	5	10.9	201	2	E83780	hypothetical prote	975	5	10.9	214	2	A35063	hypothetical prote
903	5	10.9	202	1	A45029	brefeldin A resist	976	5	10.9	215	2	S64108	hypothetical prote
904	5	10.9	202	2	G81897	probable phage ant	977	5	10.9	215	2	T01349	hypothetical prote
905	5	10.9	202	2	AG1280	1-acylglycerol-3-p	978	5	10.9	215	2	F90312	conserved hypotet

979 5 10.9 215 2 A11663 amino acid (glutam
980 5 10.9 216 1 S42716 calcineurin regula
981 5 10.9 216 2 E9128 ribosomal protein
982 5 10.9 216 2 A60331 pilin precursor -
983 5 10.9 216 2 A86633 conserved hypothet
984 5 10.9 216 2 T30657 hypothetical prote
985 5 10.9 216 2 B2612 hypothetical prote
986 5 10.9 216 2 A82548 hypothetical prote
987 5 10.9 217 1 H71208 hypothetical prote
988 5 10.9 217 2 H71203 hypothetical prote
989 5 10.9 217 2 H70123 hypothetical prote
990 5 10.9 217 2 E95370 hypothetical prote
991 5 10.9 218 2 S37658 drought-induced pr
992 5 10.9 218 2 A64482 hypothetical prote
993 5 10.9 218 2 A69032 conserved hypothet
994 5 10.9 218 2 A70964 probable deda prot
995 5 10.9 218 2 T23318 hypothetical prote
996 5 10.9 218 2 C59712 mutants block spor
997 5 10.9 218 2 B2470 hypothetical prote
998 5 10.9 219 2 A82888 d-ribulose-5-phosp
999 5 10.9 219 2 E97183 uncharacterized se
1000 5 10.9 220 2 S68211 Ig heavy chain (Ma

ALIGNMENTS

RESULT 1
D70128
conserved hypothetical protein BB0228 - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: D70128
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kervage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A:Reference number: A70100; MDID:98065943
A:Accession: D70128
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-971 <KLE>
A:Cross-references: GB:AE001133; GB:AE000783; NID:g2688120; PIDN:AAC66621.1; PID:g268812
A:Experimental source: strain B31

Query Match 19.6%; Score 9; DB 2; Length 971;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TSLKNDK 12
|||||||
Db 653 TSLKNDK 661

RESULT 2
G90517
hypothetical protein MYP0_0470 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: G90517
R:Chambaud, I.; Hellig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;
Nucleic Acids Res 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm
A:Reference number: A99512; MDID:21267165; PMID:11353084
A:Accession: G90517
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-88 <KUR>
A:Cross-references: GB:AL445566; PID:g14089460; PIDN:CAC13220.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:

A:Gene: MYP0_0470
A:Genetic code: Scc3
Query Match 15.2%; Score 7; DB 2; Length 88;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 KTLKLN 9
|||||||
Db 20 KTLKLN 26
RESULT 3
S46737
hypothetical protein YHR038w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein H8179.10
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-2001
C:Accession: S46737
R:Du, Z.
submitted to the EMBL Data Library, May 1994
A:Description: The sequence of S. cerevisiae cosmid 8179.
A:Reference number: S46732
A:Accession: S46737
A:Molecule type: DNA
A:Residues: 1-230 <DUZ>
A:Cross-references: EMBL:U00062; NID:g488162; PID:g488171; GSPDB:GN000008; MIPS:YHR038
C:Genetics:
A:Gene: MIPS:YHR038w
A:Map position: 8R

Query Match 15.2%; Score 7; DB 2; Length 230;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 34 RLNRSP 40
|||||||
Db 17 RLNRSP 23

RESULT 4
F83488
probable ATP-binding component of ABC transporter PA1256 [imported] - Pseudomonas aer
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: F83488
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: A82950; MDID:20437337
A:Accession: F83488
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <STO>
A:Cross-references: GB:AE004555; GB:AE004091; NID:g9947183; PIDN:AAG04645.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA1256

Query Match 15.2%; Score 7; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 IIGSGS 28
|||||||
Db 32 IIGSGS 38

RESULT 5
H95873
probable amino acid ABC transporter ATP-binding protein SMB20266 [imported] - Sinorhi

C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: H95873
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan-
 Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
 A:Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo-
 A:Reference number: A95842; MUID:21396508; PMID:11481431
 A:Accession: H95873
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <KUR>
 A:Cross-references: GB:AL501985; PIDN:CAC48656.1; PID:gl15140128; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymb
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wellis, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96059; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMB20266
 A:Genome: plasmid

Query Match 15.2% Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
 DB 32 IIGSGS 38

RESULT 6

ABC transporter - Aquifex aeolicus
 C:Species: Agrobacterium tumefaciens
 C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
 C:Accession: G98155
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum-
 A:Reference number: A97359; PMID:11743194
 A:Accession: G98155
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <KUR>
 A:Cross-references: GB:AE007870; PIDN:AAK88769.1; PID:gl15158515; GSPDB:GN00170
 C:Genetics:
 A:Gene: AGR_L_402
 A:Map position: linear chromosome

Query Match 15.2% Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
 DB 32 IIGSGS 38

RESULT 7

Hypothetical protein Atu4681 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AE3132
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan
 ; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam-
 ster, E.W.
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AE3132
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-240 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AAU45475.1; PID:gl17743181; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu4681
 A:Map position: linear chromosome

Query Match 15.2% Score 7; DB 2; Length 240;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
 DB 32 IIGSGS 38

RESULT 8

ABC transporter - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Feb-2001
 C:Accession: B70483
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
 V.

Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A:Reference number: A70300; MUID:98196666
 A:Accession: B70483

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-253 <AQF>
 A:Cross-references: GB:AE000774; NID:G2984324; PIDN:AAAC07853.1; PID:G2984332; GB:AE00

A:Experimental source: strain VF5
 C:Genetics:

A:Gene: abct12
 C:Superfamily: inner membrane protein malK; ATP-binding cassette homology
 C:Keywords: ATP; nucleotide binding; P-loop
 F:20-218/Domain: ATP-binding cassette homology <ABC>
 F:37-44/Region: nucleotide-binding motif A (P-loop)

Query Match 15.2% Score 7; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
 DB 35 IIGSGS 41

RESULT 9

stem cell factor short form precursor - quail
 C:Species: Coturnix coturnix (quail)
 C:Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
 C:Accession: S70367

R:Petitje, J.N.; Kulik, M.J.
 Biochim. Biophys. Acta 1307, 149-151, 1996
 A:Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell
 A:Reference number: S70366; MUID:96283808
 A:Accession: S70367

A:Molecule type: mRNA
 A:Residues: 1-253 <PET>

A:Cross-references: EMBL:U43079; NID:gl150877; PIDN:AAAC59934.1; PID:gl150878
 C:Superfamily: mouse mast cell growth factor
 C:Keywords: growth factor; transmembrane protein

F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-253/Product: stem cell factor short form #status predicted <MAT>
 F:192-216/Domain: transmembrane #status predicted <TM>

Query Match 15.2%; Score 7; DB 2; Length 253;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 FFRLFN 38
 Db 144 FFRLFN 150

RESULT 10

S70366
 stem cell factor long form precursor - quail
 C:Species: Coturnix coturnix (quail)
 C>Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
 C:Accession: S70366
 R:Petit, J.N.; Kulik, M.J.
 Biochim. Biophys. Acta 1307, 149-151, 1996
 A:Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell factor
 A:Reference number: S70366; MUID:96283808
 A:Accession: S70366
 A:Molecule type: mRNA
 A:Residues: 1-287 <PE>
 A:Cross-references: EMBL:U43078; NID:q1150875; PIDN:AAC59933.1; PID:q1150876
 C:Superfamily: mouse mast cell growth factor
 C:Keywords: growth factor; transmembrane protein
 F:1-25/Domain: signal sequence #status predicted <SIG>
 F:26-287/Product: stem cell factor long form #status predicted <MAT>
 F:226-250/Domain: transmembrane #status predicted <TM>

Query Match 15.2%; Score 7; DB 2; Length 287;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 FFRLFN 38
 Db 144 FFRLFN 150

RESULT 11

T48633
 hypothetical protein TI5N1.90 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48633
 R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224493
 A:Accession: T48633
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-370 <BEV>
 A:Cross-references: EMBL:AL163792
 A:Experimental source: cultivar Columbia; BAC clone TI5N1
 C:Genetics:
 A:Map position: 5
 A:Introns: 150/2; 232/1; 307/1
 A:Note: TI5N1.90

Query Match 15.2%; Score 7; DB 2; Length 370;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 GSGSLST 31
 Db 118 GSGSLST 124

RESULT 12

155604

platelet glycoprotein Ib beta chain precursor, endothelial splice form - human
 N:Alternate names: membrane glycoprotein Ib beta chain (GP1b)
 C:Species: Homo sapiens (man)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
 C:Accession: I55604
 R:Kelly, M.D.; Essex, D.W.; Shapiro, S.S.; Meloni, F.J.; Druck, T.; Huebner, K.; Konk
 J. Clin. Invest. 93, 2417-2424, 1994
 A:Title: Complementary DNA cloning of the alternatively expressed endothelial cell gl
 A:Reference number: I55604; MUID:94259799
 A:Accession: I55604
 A:Status: translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-411 <KEL>
 A:Cross-references: GB:L20860; NID:g438627; PIDN:AAA20398.1; PID:g517494

C:Genetics:

A:Gene: GDB:GPIIB
 A:Cross-references: GDB:I28731; OMIM:138720
 A:Map position: 22q11.21-22q11.23
 A:Introns: 209/1
 C:Complex: heterodimer with platelet glycoprotein Ib alpha chain (PIR:NBUHTA)
 C:Superfamily: platelet glycoprotein Ib beta chain; leucine-rich alpha-2-glycoprotein
 C:Keywords: alternative splicing; blood coagulation; duplication; glycoprotein; heter

F:224-249/Domain: proteoglycan amino-terminal homology <PAH>
 F:261-284/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
 F:292-338/Domain: proteoglycan carboxyl-terminal homology <PCH>
 F:353-377/Domain: transmembrane #status predicted <TRM>
 F:271/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:396/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 15.2%; Score 7; DB 1; Length 411;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GSGSLS 30
 Db 45 GSGSLS 51

RESULT 13

AC1112
 PTS beta-glucoside-specific enzyme IIC component homolog lmo0298 [imported] - Listeri
 C:Species: Listeria monocytogenes
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 31-Dec-2001
 C:Accession: AC1112
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
 ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1112
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00825.1; PID:g16409663; GSPDB:GNO0177
 A:Experimental source: strain B3D-e
 C:Genetics:
 A:Gene: lmo0298
 C:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-dep

Query Match 15.2%; Score 7; DB 2; Length 437;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 ITGGSGS 28
 Db 293 ITGGSGS 299

RESULT 14

AG1473

PTS beta-glucoside-specific enzyme IIC component homolog lin0326 [imported] - Listeria
 C:Species: Listeria innocua
 C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 31-Dec-2001
 C:Accession: AG1473
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
 D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
 Science 294, 849-852, 2001
 A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
 ok, C.; Schluter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1473
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-437 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC95559.1; PID:gi6412755; GSPDB:GN00178
 A:Experimental source: strain Clp11262
 C:Genetics:
 A:Gene: lin0326
 C:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 15.2%; Score 7; DB 2; Length 437;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28

Db 293 IIGSGS 299

RESULT 15

AD2284
 hypothetical protein alr3827 [imported] - Anabaena sp. (strain PCC 7120)
 C:Species: Anabaena sp.
 A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002
 C:Accession: AD2284
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
 DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: AD2284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-472 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BA075526.1; PID:gi17132961; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3827
 C:Superfamily: synchocystis hypothetical protein sll0827

Query Match 15.2%; Score 7; DB 2; Length 472;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 SFTQALG 45

Db 408 SFTQALG 414

Search completed: November 5, 2002, 11:06:57
 Job time : 69.8507 secs

GenCore version 5.1.3

Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:58:02 ; Search time 12.3582 Seconds
(without alignments)
144.123 Million cell updates/sec

Title: US-09-833-017B-2

Perfect score: 46

Sequence: 1 MKKTLKNDKFKIKTELE.....GSLSTFFRLNRSTQALGK 46

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	19.6	971	1 Y228_BORBU	O51246 borrelia bu
2	7	15.2	230	1 FIL1_YEAST	P38771 saccharomyc
3	7	15.2	500	1 ABFI_TRE	Q92455 trichoderma
4	7	15.2	500	1 XYLI_TRIKO	P48792 trichoderma
5	7	15.2	558	1 RELB_MOUSE	O04863 mus musculu
6	7	15.2	579	1 RELB_HUMAN	Q01201 homo sapien
7	7	15.2	591	1 YN48_YEAST	P43846 saccharomyc
8	7	15.2	788	1 UGSS_SOLITU	Q43847 solanum tub
9	7	15.2	791	1 K6PP_RABIT	P47859 oryctolagus
10	6	13.0	88	1 PSD8_DICDI	P02889 dictyosteli
11	6	13.0	110	1 CYC6_PORPU	P51200 porphyra pu
12	6	13.0	114	1 HV3B_HUMAN	P01763 homo sapien
13	6	13.0	122	1 HV3A_HUMAN	P01762 homo sapien
14	6	13.0	142	1 RADC_COXBU	O85403 coxiella bu
15	6	13.0	155	1 Y805_AQUAE	O66989 aquifex aeo
16	6	13.0	169	1 CCAC_CAVPO	O35505 cavia porce
17	6	13.0	177	1 CCAC_CHICK	O73707 gallus gall
18	6	13.0	214	1 Y540_AQUAE	O66819 aquifex aeo
19	6	13.0	216	1 CAT1_STAAR	P00485 staphylococ
20	6	13.0	219	1 PSB9_HUMAN	P28065 homo sapien
21	6	13.0	219	1 PSB9_MOUSE	P28076 mus musculu
22	6	13.0	219	1 PSB9_MUSDU	O35521 mus dunni
23	6	13.0	219	1 PSB9_MUSPL	O35523 mus platyth
24	6	13.0	219	1 PSB9_MUSSI	O35524 mus spicile
25	6	13.0	219	1 PSB9_RAT	P28077 rattus norv
26	6	13.0	226	1 PSB6_SCHPO	O43063 schizosacch
27	6	13.0	234	1 PSB6_TOHAC	P93395 nicotiana t
28	6	13.0	237	1 PSB6_RAT	P28073 rattus norv
29	6	13.0	237	1 Y417_METJA	Q57860 methanococc
30	6	13.0	238	1 PSB6_MOUSE	Q60692 mus musculu
31	6	13.0	239	1 PSB6_HUMAN	P28072 homo sapien
32	6	13.0	257	1 Y125_PYRHO	O57865 pyrococcus
33	6	13.0	257	1 YN64_PYRAB	Q9V2f1 pyrococcus

34	6	13.0	273	1 DAPF_HELPJ	Q921r5 helicobacte
35	6	13.0	273	1 DAPF_HELPJ	O23290 helicobacte
36	6	13.0	276	1 OCTT_AGRTR	P35121 agrobacteri
37	6	13.0	281	1 CCAD_MOUSE	O99246 mus musculu
38	6	13.0	281	1 CCAS_CHICK	O42398 gallus gall
39	6	13.0	285	1 MTAP_CAEEL	Q09438 caenorhabdi
40	6	13.0	287	1 SCF_CHICK	O09108 gallus gall
41	6	13.0	296	1 CEBB_MOUSE	P28033 mus musculu
42	6	13.0	297	1 CEBB_RAT	P21272 rattus norv
43	6	13.0	309	1 HLCA_STAAM	P31714 staphylococ
44	6	13.0	311	1 OPFC_HAEIN	P45053 haemophilus
45	6	13.0	314	1 FMT_BUCAL	P57564 buchnera ap
46	6	13.0	316	1 Y441_CHLPN	O928a2 chlamydia p
47	6	13.0	317	1 FATC_VIBAN	P77737 vibrio angu
48	6	13.0	329	1 ISPB_HAEIN	P44916 haemophilus
49	6	13.0	342	1 FLGI_HELPJ	Q92mi9 helicobacte
50	6	13.0	342	1 FLGI_HELPJ	O25028 helicobacte
51	6	13.0	344	1 ALP_MYCTU	O06313 mycobacteri
52	6	13.0	345	1 CATD_PIG	P00795 sus scrofa
53	6	13.0	345	1 CEBB_HUMAN	P17676 homo sapien
54	6	13.0	348	1 CEBB_BOVIN	O02755 bos taurus
55	6	13.0	349	1 YGAL_YEAST	P53199 s putative
56	6	13.0	354	1 FXFL_HUMAN	Q12946 homo sapien
57	6	13.0	365	1 CATD_SHEEP	Q9mzs8 ovine aries
58	6	13.0	372	1 INVE_SALTY	P35671 salmonella
59	6	13.0	373	1 ACT6_DIPDE	P53459 diphylobot
60	6	13.0	375	1 TISX_MYCPN	P75159 mycoplasma
61	6	13.0	387	1 ASPP_AEGAE	O03168 aedes aegypt
62	6	13.0	390	1 CATD_BOVIN	P80209 bos taurus
63	6	13.0	393	1 CYB_VENIN	O48334 venturia in
64	6	13.0	393	1 YBZJ_CAEEL	P91119 caenorhabdi
65	6	13.0	396	1 CARP_NEUCR	Q01294 neurospora
66	6	13.0	405	1 WCAC_ECOLI	P71237 escherichia
67	6	13.0	407	1 CATD_RAT	P24268 rattus norv
68	6	13.0	410	1 CATD_MOUSE	P18242 mus musculu
69	6	13.0	412	1 CATD_HUMAN	P07339 homo sapien
70	6	13.0	414	1 PYRC_THEAC	O9hno0 thermoplas
71	6	13.0	417	1 DSRA_ARCFU	O59109 archaeoglob
72	6	13.0	442	1 SYN_HELPJ	Q92k27 helicobacte
73	6	13.0	442	1 APRA_PSEAB	P56455 helicobacte
74	6	13.0	479	1 SYH_HELPJ	Q30033 pseudomonas
75	6	13.0	484	1 R101_YEAST	Q12136 saccharomyc
76	6	13.0	494	1 STN1_YEAST	P38960 saccharomyc
77	6	13.0	497	1 UXAA_BACSU	O34673 bacillus su
78	6	13.0	499	1 ABFB_ASPNG	P42255 aspergillus
79	6	13.0	499	1 NU4C_MARPO	O06263 marichantia
80	6	13.0	506	1 Y619_METJA	Q58036 methanococc
81	6	13.0	518	1 LEU1_BACSU	P94565 bacillus su
82	6	13.0	521	1 GAG_HV2D2	P15832 human immun
83	6	13.0	538	1 RO60_XENIA	P42700 xenopus lae
84	6	13.0	550	1 MP10_XENIA	P30308 xenopus lae
85	6	13.0	550	1 MP11_XENIA	P30309 xenopus lae
86	6	13.0	551	1 CH60_BUCAP	O59177 buchnera ap
87	6	13.0	561	1 YEHU_ECOLI	P33357 escherichia
88	6	13.0	569	1 CH63_HELVI	P25430 heliothis v
89	6	13.0	572	1 MP13_XENIA	P30311 xenopus lae
90	6	13.0	573	1 CH60_CRIGR	P18687 cricetus
91	6	13.0	573	1 CH60_HUMAN	P10809 homo sapien
92	6	13.0	573	1 CH60_MOUSE	P19226 mus musculu
93	6	13.0	599	1 MP12_XENIA	P30310 xenopus lae
94	6	13.0	617	1 SYP_SYNY3	P73942 synecocyst
95	6	13.0	620	1 S6A6_HUMAN	P31641 homo sapien
96	6	13.0	634	1 SL56_RAT	O70247 rattus norv
97	6	13.0	635	1 SL56_HUMAN	O9y289 homo sapien
98	6	13.0	636	1 SL56_RABIT	Q9xt77 oryctolagus
99	6	13.0	637	1 MUTA_PROPR	P11652 propionibac
100	6	13.0	639	1 ACTA_LISMO	P33379 listeria mo
101	6	13.0	664	1 UVRB_THEMA	Q9x282 thermotoga
102	6	13.0	699	1 YQQA_CAEEL	Q09299 caenorhabdi
103	6	13.0	703	1 ML51_HUMAN	O15234 homo sapien
104	6	13.0	717	1 AMYM_BACST	P19531 bacillus st
105	6	13.0	735	1 CNGL_CHICK	Q90805 gallus gall
106	6	13.0	759	1 PMT6_YEAST	P42934 saccharomyc

107	6	13.0	830	1	GYRA_BUCAL	P57277	buchnera ap	180	5	10.9	102	1	VAE5_YEAST	P39725	saccharomyc
108	6	13.0	833	1	VIRA_AGR5	P18540	agrobacteri	181	5	10.9	103	1	CYT1_MOUSE	P35173	mus musculus
109	6	13.0	841	1	YAGX_ECOLI	P77802	escherichia	182	5	10.9	104	1	YGI9_YEAST	P53069	saccharomyc
110	6	13.0	879	1	YEO7_MYCPN	P75377	mycoplasma	183	5	10.9	106	1	Y116_ADE02	P03267	human adeno
111	6	13.0	887	1	HM0H_CRIGR	P00347	cricketul	184	5	10.9	106	1	Y488_ARCFU	P02962	archaeoglob
112	6	13.0	887	1	HM0H_MESAU	P09610	mesocricetu	185	5	10.9	108	1	YDBL_ECOLI	P71770	escherichia
113	6	13.0	887	1	HM0H_RAT	P51639	rattus norv	186	5	10.9	108	1	YB89_MYCTU	P71770	mycobacteri
114	6	13.0	888	1	HM0H_HUMAN	P04035	homo sapien	187	5	10.9	109	1	FEL2_FELXAL	P30440	felis silve
115	6	13.0	888	1	HM0H_RABIT	P09512	oryctolagus	188	5	10.9	109	1	PER_MUSDO	P05440	loxocera al
116	6	13.0	970	1	Y087_BUCAL	P57189	buchnera ap	189	5	10.9	109	1	PER_MUSDO	P05440	musca domes
117	6	13.0	1049	1	PM02_SCHPO	P03824	schizosacch	190	5	10.9	109	1	PER_RHACO	P02631	rhagoletis
118	6	13.0	1075	1	PM02_SCHPO	P03824	schizosacch	191	5	10.9	109	1	PER_SYPRI	P02631	syntherisma
119	6	13.0	1126	1	PM02_SCHPO	P03824	schizosacch	192	5	10.9	110	1	PER_SYPRI	P02631	syntherisma
120	6	13.0	1126	1	PM02_SCHPO	P03824	schizosacch	193	5	10.9	110	1	PER_SYPRI	P02631	syntherisma
121	6	13.0	1126	1	PM02_SCHPO	P03824	schizosacch	194	5	10.9	110	1	PER_SYPRI	P02631	syntherisma
122	6	13.0	1373	1	RPOB_RICCN	P09441	rickettsia	195	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
123	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	196	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
124	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	197	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
125	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	198	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
126	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	199	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
127	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	200	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
128	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	201	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
129	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	202	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
130	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	203	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
131	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	204	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
132	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	205	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
133	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	206	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
134	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	207	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
135	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	208	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
136	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	209	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
137	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	210	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
138	6	13.0	1374	1	RPOB_RICCN	P09441	rickettsia	211	5	10.9	113	1	YV11_FPVLL	P06456	avian papil
139	5	10.9	26	1	YFHA_KLEPN	P07170	klebsiella	212	5	10.9	124	1	PER_HIRPI	P10773	halobacteri
140	5	10.9	27	1	YFHA_KLEPN	P07170	klebsiella	213	5	10.9	124	1	PER_HIRPI	P10773	halobacteri
141	5	10.9	33	1	ACT1_DICVI	P07170	klebsiella	214	5	10.9	125	1	PER_HIRPI	P10773	halobacteri
142	5	10.9	40	1	ACT1_DICVI	P07170	klebsiella	215	5	10.9	125	1	PER_HIRPI	P10773	halobacteri
143	5	10.9	45	1	CYC6_PROHO	P19007	oryctolagus	216	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
144	5	10.9	55	1	A70A_DROMA	P19007	oryctolagus	217	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
145	5	10.9	56	1	RK33_GUTH	P19007	oryctolagus	218	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
146	5	10.9	60	1	LANP_STAEP	P19007	oryctolagus	219	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
147	5	10.9	62	1	Y773_ARCFU	P03294	drosophila	220	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
148	5	10.9	63	1	PER_DROMS	P03294	drosophila	221	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
149	5	10.9	63	1	PER_DROMS	P03294	drosophila	222	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
150	5	10.9	65	1	PER_DROMS	P03294	drosophila	223	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
151	5	10.9	65	1	PER_DROMS	P03294	drosophila	224	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
152	5	10.9	66	1	PER_DROMS	P03294	drosophila	225	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
153	5	10.9	67	1	PER_DROMS	P03294	drosophila	226	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
154	5	10.9	67	1	PER_DROMS	P03294	drosophila	227	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
155	5	10.9	69	1	MCBA_ECOLI	P05834	escherichia	228	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
156	5	10.9	70	1	YORA_TTV1	P19285	thermoprote	229	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
157	5	10.9	82	1	Y52A_METJA	P191309	methanococc	230	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
158	5	10.9	83	1	Y9KD_BPF2	P191309	methanococc	231	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
159	5	10.9	86	1	PER_DROMS	P03294	drosophila	232	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
160	5	10.9	87	1	CYSD_CHLLT	P20958	chlorobium	233	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
161	5	10.9	87	1	IM13_ARATH	P09441	rickettsia	234	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
162	5	10.9	87	1	Y794_METJA	P09441	rickettsia	235	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
163	5	10.9	88	1	GP42_BPSPI	P08325	treponema p	236	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
164	5	10.9	88	1	PER_DROTE	P08325	treponema p	237	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
165	5	10.9	88	1	SL08_MOUSE	P27005	mus musculus	238	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
166	5	10.9	89	1	SLTB_BPH30	P08027	bacterioph	239	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
167	5	10.9	90	1	YR17_GUTH	P046903	guillardi	240	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
168	5	10.9	91	1	YR17_GUTH	P046903	guillardi	241	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
169	5	10.9	93	1	YR17_GUTH	P046903	guillardi	242	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
170	5	10.9	93	1	CBIN_SALTY	P05595	salmonella	243	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
171	5	10.9	93	1	RS19_HELPV	P05595	salmonella	244	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
172	5	10.9	93	1	RS19_HELPV	P05595	salmonella	245	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
173	5	10.9	94	1	ES46_MYCTU	P56026	helicobacte	246	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
174	5	10.9	94	1	VA15_VACCC	P07165	mycobacteri	247	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
175	5	10.9	94	1	VA15_VACCC	P07165	mycobacteri	248	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
176	5	10.9	95	1	VA15_VACCC	P07165	mycobacteri	249	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
177	5	10.9	95	1	VA15_VACCC	P07165	mycobacteri	250	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
178	5	10.9	96	1	HO1I_BPA18	P03940	variola vir	251	5	10.9	126	1	PER_HIRPI	P10773	halobacteri
179	5	10.9	96	1	HO1I_BPA18	P03940	variola vir	252	5	10.9	126	1	PER_HIRPI	P10773	halobacteri

253 5 10.9 160 1 RT2C_ACTPL
254 5 10.9 161 1 ACT_PINGO
255 5 10.9 161 1 Y311_RICPR
256 5 10.9 161 1 Y447_METJA
257 5 10.9 164 1 Y07E_BP4
258 5 10.9 164 1 YPOC_BACSU
259 5 10.9 166 1 K2C5_BOVIN
260 5 10.9 167 1 LSPA_CHLMU
261 5 10.9 167 1 NADM_METJA
262 5 10.9 167 1 TCTP_LUMRU
263 5 10.9 169 1 CALB_HUMAN
264 5 10.9 169 1 CALB_MOUSE
265 5 10.9 169 1 CME_BRAJA
266 5 10.9 169 1 RST4_AVEA
267 5 10.9 171 1 S3AB_BACSU
268 5 10.9 172 1 ACT3_LYTP1
269 5 10.9 172 1 ACTM_LYTP1
270 5 10.9 172 1 SSB_BACSU
271 5 10.9 172 1 VP19_TBSVB
272 5 10.9 172 1 VP19_TBSVT
273 5 10.9 173 1 MBP_CHICK
274 5 10.9 173 1 PRLA_HORVU
275 5 10.9 173 1 PRIC_HORVU
276 5 10.9 173 1 TLP_WHEAT
277 5 10.9 173 1 YFHC_HAEIN
278 5 10.9 174 1 GRPE_METTH
279 5 10.9 175 1 CALC_RAT
280 5 10.9 175 1 RL6B_YEAST
281 5 10.9 177 1 ATPD_BUCAP
282 5 10.9 178 1 AIL_YEREN
283 5 10.9 178 1 CALC_MOUSE
284 5 10.9 178 1 KITH_MYXVA
285 5 10.9 178 1 YC96_HAEIN
286 5 10.9 179 1 PGSG_RAT
287 5 10.9 179 1 RF5_EUGGR
288 5 10.9 180 1 ARF1_PLAFO
289 5 10.9 180 1 ARF_PLAFA
290 5 10.9 181 1 APT2_YEAST
291 5 10.9 181 1 FANC_ECOLI
292 5 10.9 181 1 Y65L_HORVU
293 5 10.9 182 1 K2CA_BOVIN
294 5 10.9 183 1 DIN1_RAPSA
295 5 10.9 183 1 GRP2_ORYSA
296 5 10.9 186 1 DYR_HSV57
297 5 10.9 187 1 DYR_HSV5A
298 5 10.9 188 1 COME_METJA
299 5 10.9 188 1 EEP_PASMU
300 5 10.9 190 1 TRPG_CYAPA
301 5 10.9 191 1 KIP1_HUMAN
302 5 10.9 191 1 KIP1_MOUSE
303 5 10.9 191 1 KIP1_RAT
304 5 10.9 191 1 KITH_STRGC
305 5 10.9 192 1 LPCA_HELPJ
306 5 10.9 192 1 LPCA_HELPY
307 5 10.9 192 1 MOBA_PYROH
308 5 10.9 193 1 DEOC_AERPE
309 5 10.9 193 1 YBF9_YEAST
310 5 10.9 193 1 YHMF_HAEIN
311 5 10.9 195 1 KITH_BACNA
312 5 10.9 196 1 AEQ2_AEQVI
313 5 10.9 196 1 CRF_HUMAN
314 5 10.9 197 1 RUVA_CLOHI
315 5 10.9 197 1 Y979_METJA
316 5 10.9 198 1 PDX2_MOUSE
317 5 10.9 198 1 PDX2_RAT
318 5 10.9 199 1 RSBX_BACSU
319 5 10.9 199 1 TDX2_BRUMA
320 5 10.9 201 1 TRI_PONLE
321 5 10.9 202 1 P25_SCHPO
322 5 10.9 203 1 KTHV_CHLMU
323 5 10.9 203 1 VP10_BPPRD
324 5 10.9 203 1 Y260_METJA
325 5 10.9 204 1 CAHB_MOUSE

326 5 10.9 205 1 FADD_MOUSE
327 5 10.9 205 1 URK_BORBU
328 5 10.9 207 1 3MGR_LISMO
329 5 10.9 207 1 IM22_YEAST
330 5 10.9 207 1 RID2_BUCAI
331 5 10.9 210 1 CAT4_MORMO
332 5 10.9 211 1 HFA5_HAEIN
333 5 10.9 211 1 URK_BACSU
334 5 10.9 211 1 HFA3_HAEIN
335 5 10.9 212 1 HFA3_HAEIN
336 5 10.9 212 1 RBL7_HUMAN
337 5 10.9 213 1 DYR_HSVSC
338 5 10.9 213 1 HFA2_HAEIN
339 5 10.9 213 1 HFA4_HAEIN
340 5 10.9 213 1 PSMB_ARCFU
341 5 10.9 214 1 ACH2_LONAC
342 5 10.9 214 1 YL31_HALMA
343 5 10.9 214 1 YQ37_MYCLE
344 5 10.9 215 1 YGK1_YEAST
345 5 10.9 215 1 HFA1_HAEIN
346 5 10.9 216 1 RSS_METTH
347 5 10.9 216 1 S3AH_BACSU
348 5 10.9 218 1 YC35_METTH
349 5 10.9 218 1 Y558_METJA
350 5 10.9 218 1 YQ37_MYCTU
351 5 10.9 221 1 YD1E_SCHPO
352 5 10.9 222 1 FTSQ_CORGL
353 5 10.9 222 1 RISB_SPIOL
354 5 10.9 222 1 COAT_TRVCA
355 5 10.9 224 1 GSHU_CABEL
356 5 10.9 224 1 PSMB_METJA
357 5 10.9 224 1 Y088_MYCTU
358 5 10.9 224 1 YXEN_BACSU
359 5 10.9 227 1 DJB8_MOUSE
360 5 10.9 228 1 PCRB_BACSU
361 5 10.9 229 1 GLC8_YEAST
362 5 10.9 229 1 XYN1_TRIRE
363 5 10.9 233 1 DEOD_TREPA
364 5 10.9 233 1 YSR3_CABEL
365 5 10.9 234 1 EMB2_CAVFO
366 5 10.9 236 1 ABME_HUMAN
367 5 10.9 236 1 Y545_RICPR
368 5 10.9 237 1 CD63_MOUSE
369 5 10.9 237 1 CD63_RAT
370 5 10.9 237 1 LAPT_PASHA
371 5 10.9 238 1 YB92_YEAST
372 5 10.9 238 1 YPE2_SHORU
373 5 10.9 239 1 ARTI_HAEIN
374 5 10.9 239 1 RP35_BACTR
375 5 10.9 239 1 VIB1_AGR79
376 5 10.9 240 1 H12_VOLCA
377 5 10.9 240 1 LEF5_GVCL
378 5 10.9 240 1 YMFC_HAEIN
379 5 10.9 241 1 Y293_MYCPN
380 5 10.9 243 1 GUB_BACLI
381 5 10.9 243 1 TONE_PSEPU
382 5 10.9 244 1 Y293_MYGGE
383 5 10.9 246 1 VGLU_HSV7J
384 5 10.9 247 1 IF6_ARATH
385 5 10.9 247 1 RL7_ARATH
386 5 10.9 248 1 GRL1_RAT
387 5 10.9 248 1 NUKC_MAIZE
388 5 10.9 250 1 MTRD_METMA
389 5 10.9 250 1 YCH6_YEAST
390 5 10.9 251 1 TX20_HUMAN
391 5 10.9 252 1 PMM_CANAL
392 5 10.9 252 1 TRPA_AERPE
393 5 10.9 252 1 TRY1_DROME
394 5 10.9 252 1 YC88_METJA
395 5 10.9 252 1 YKC6_CAEEL
396 5 10.9 253 1 NIKD_ECOLI
397 5 10.9 254 1 MUR1_AQUAE
398 5 10.9 254 1 RACD_DICDI

Q61160 mus musculus
Q59190 borrelia bu
Q52d89 listeria in
Q58621 listeria mo
Q12328 saccharomyc
P57534 buchnera ap
P50869 morganella
Q45930 haemophilus
Q32033 bacillus su
P45988 haemophilus
Q9h0k7 homo sapien
P22573 herpesvirus
P14212 haemophilus
P45989 haemophilus
Q9p996 archaeoglob
P23605 ionomia ach
P20571 haloarcula
Q49642 mycobacteri
P53144 saccharomyc
Q03846 haemophilus
Q26131 methanobact
P49795 bacillus su
Q27303 methanobact
Q38853 methanococc
P71936 mycobacteri
Q10244 schizosacch
P94336 corynebacte
Q9xh32 spinacia ol
P05070 tobacco rat
Q95003 caenorhabdi
Q38634 methanococc
Q10885 mycobacteri
P34953 bacillus su
Q9qy17 mus musculu
Q34790 bacillus su
P41818 saccharomyc
P36218 trichoderma
Q83716 treponema p
Q9951 caenorhabdi
P35709 cavia porce
P41238 homo sapien
Q92d05 rickettsia
P41731 mus musculu
P28648 rattus norv
P49618 pasteurella
P38331 saccharomyc
P23139 rhodospiril
P45091 haemophilus
P26763 bacillus th
P05330 agrobacteri
Q08865 volvox cart
P41727 cryptophleb
P44827 haemophilus
P75367 mycoplasma
Q27051 bacillus li
Q05613 pseudomonas
P47535 mycoplasma
P32509 human herpe
Q22290 arabidopsi
Q42208 arabidopsi
Q06605 rattus norv
P06670 zea mays (m
P80653 methanosarc
P25659 saccharomyc
Q9umr3 homo sapien
P31353 candida alb
Q9y8t3 aeropyrum p
P52905 drosophila
Q58684 methanococc
P41997 caenorhabdi
P33593 escherichia
Q66662 aquifex aeo
P34150 dictyostell

545	5	10.9	327	1	GRE3_YEAST	P38715	saccharomyc	618	5	10.9	348	1	VNAT_PIIHA	P36355	human parai
546	5	10.9	328	1	CAHE_HUMAN	O75493	homo sapien	619	5	10.9	348	1	VNAT_PIIHC	Q01427	human parai
547	5	10.9	328	1	CAHE_SHEEP	O95203	ovis aries	620	5	10.9	348	1	VNAT_SENDS	P27567	sendai viru
548	5	10.9	330	1	YCDU_ECOLI	P75910	escherichia	621	5	10.9	348	1	VNAT_SENDF	P17748	sendai viru
549	5	10.9	330	1	GCAA_MOUSE	P01863	mus musculus	622	5	10.9	348	1	VNAT_SENDF	P03426	sendai viru
550	5	10.9	330	1	YDEZ_ECOLI	P77651	escherichia	623	5	10.9	348	1	VNAT_SENDF	P06446	sendai viru
551	5	10.9	331	1	ACT_COSCS	P30161	costaria co	624	5	10.9	349	1	GPDA_DROAE	Q27556	drosophila
552	5	10.9	331	1	APL_HUMAN	P56412	homo sapien	625	5	10.9	349	1	GPDA_DROEZ	Q27567	drosophila
553	5	10.9	331	1	APL_PIG	P54332	sus scrofa	626	5	10.9	349	1	ISPG_CLOPE	P58667	clostridium
554	5	10.9	331	1	PLIE_ACKBL	O93233	agkistrodon	627	5	10.9	350	1	AROG_ECOLI	P00886	escherichia
555	5	10.9	332	1	ACT1_SOLTU	P93587	solanum tub	628	5	10.9	350	1	OSTG_YEAST	P48439	saccharomyc
556	5	10.9	332	1	HME2_HUMAN	P19622	homo sapien	629	5	10.9	350	1	SUBI_SYNP7	P27366	synecococc
557	5	10.9	332	1	RNFED_BUCAI	P57216	buchnera ap	630	5	10.9	350	1	YCJQ_ECOLI	P76043	escherichia
558	5	10.9	332	1	SSUA_BACSU	P40400	bacillus su	631	5	10.9	350	1	YCXA_EUGGR	P31561	euglena gra
559	5	10.9	333	1	TAL2_YEAST	P53228	saccharomyc	632	5	10.9	351	1	VNAT_PI3B	P06166	bovine para
560	5	10.9	333	1	YZ37_SYNY3	Q55480	synecocyst	633	5	10.9	352	1	GPDA_DROVI	P07735	drosophila
561	5	10.9	335	1	GCAE_MOUSE	P01864	mus musculus	634	5	10.9	352	1	KITH_HSYMD	P17653	marek's dis
562	5	10.9	335	1	GVPN_ANAFL	P55150	anaeana fl	635	5	10.9	352	1	OCT3_MOUSE	P20263	mus musculus
563	5	10.9	336	1	ACT1_LYCES	Q9482	lycopersico	636	5	10.9	352	1	SUBI_SYNY3	Q01903	synecocyst
564	5	10.9	336	1	ACT2_LYCES	Q9483	lycopersico	637	5	10.9	353	1	ACT1_ACECL	P53491	acetabulari
565	5	10.9	336	1	ACT2_SOLTU	P93586	solanum tub	638	5	10.9	353	1	GPDA_DROPS	Q27528	drosophila
566	5	10.9	336	1	ACT2_TOBAC	P93374	nicotiana t	639	5	10.9	353	1	HIS7_BUCAI	P57203	buchnera ap
567	5	10.9	336	1	ACT3_LYCES	Q9484	lycopersico	640	5	10.9	353	1	HMV2_METJA	Q58734	methanococc
568	5	10.9	336	1	ACT4_LYCES	Q9481	lycopersico	641	5	10.9	353	1	OP28_HAEIN	Q48221	haemophilus
569	5	10.9	336	1	ACT4_TOBAC	P93372	nicotiana t	642	5	10.9	353	1	SB11_CABEL	P46506	caenorhabdi
570	5	10.9	336	1	ACT5_SOLTU	P81228	solanum tub	643	5	10.9	353	1	T2BA_BACAR	P19887	bacillus an
571	5	10.9	336	1	ACT5_TOBAC	P93371	nicotiana t	644	5	10.9	353	1	VNAT_PI3H4	P07873	human parai
572	5	10.9	336	1	ACT6_TOBAC	P93376	nicotiana t	645	5	10.9	354	1	ALF_CAMJE	P53818	campylobact
573	5	10.9	336	1	ACT7_TOBAC	P93375	nicotiana t	646	5	10.9	354	1	GLN1_ARATH	Q91vi8	arabidopsis
574	5	10.9	336	1	ACT8_SOLTU	P81229	solanum tub	647	5	10.9	354	1	PONI_MOUSE	P52430	mus musculus
575	5	10.9	336	1	ACT9_SOLTU	P93584	solanum tub	648	5	10.9	355	1	GLN1_PEA	P08282	pisum sativ
576	5	10.9	336	1	AGLF_RHIME	Q923r6	rhizobium m	649	5	10.9	355	1	GLN1_SOYBN	P24099	glycine max
577	5	10.9	336	1	GC8_MOUSE	P01866	mus musculus	650	5	10.9	355	1	GLN4_MAIZE	P38562	zea mays (m
578	5	10.9	336	1	Y05G_BPT4	P39242	bacterioph	651	5	10.9	355	1	NIFS_LACDE	P31672	lactobacill
579	5	10.9	337	1	A85A_MYCU	P58248	mycobacteri	652	5	10.9	355	1	Y816_DROME	Q9vaf0	drosophila
580	5	10.9	337	1	ACT4_SOLTU	P93585	solanum tub	653	5	10.9	356	1	DXR_CAMJE	Q9pav3	campylobact
581	5	10.9	337	1	PLSX_BAGHD	Q9ka01	bacillus ha	654	5	10.9	356	1	GLN1_ALNGL	Q04867	alnus glut
582	5	10.9	337	1	TDAB_MOUSE	Q61038	mus musculus	655	5	10.9	356	1	GLN1_LORTA	Q42899	lotus japon
583	5	10.9	337	1	VFI1_VACCP	P29888	vaccinia vi	656	5	10.9	356	1	GLN1_MEDSA	P04078	medicago sa
584	5	10.9	337	1	YJMC_BACSU	Q34736	bacillus su	657	5	10.9	356	1	GLN1_PHAVU	P04770	phaseolus v
585	5	10.9	338	1	CTE2_HUMAN	O00154	homo sapien	658	5	10.9	356	1	GLN2_VITVI	P51119	vitis vinif
586	5	10.9	338	1	CTE2_RAT	Q64559	rattus norv	659	5	10.9	356	1	GLN3_HORVU	Q06378	hordeum vul
587	5	10.9	338	1	FLIG_PSEAE	Q51464	pseudomonas	660	5	10.9	356	1	GLN3_MAIZE	P38561	zea mays (m
588	5	10.9	338	1	YBD1_YEAST	P38200	saccharomyc	661	5	10.9	356	1	GLN3_ORYSA	P14656	oryza sativ
589	5	10.9	339	1	YAF9_SCHPO	Q09862	schizosacch	662	5	10.9	356	1	GLNA_NICPL	P12424	nicotiana p
590	5	10.9	340	1	ALF_STRGB	Q92em7	streptomyce	663	5	10.9	356	1	GLNA_VIGAC	P32389	vigna acon
591	5	10.9	340	1	THPS_SULAC	P17118	bulbobotus	664	5	10.9	357	1	ACT1_OXYFA	P53503	oxytricha f
592	5	10.9	341	1	BMPS_BORAF	Q31284	borrelia af	665	5	10.9	357	1	ACT2_OXYFA	P02583	oxytricha f
593	5	10.9	341	1	VPI3_GFLV	P17768	grapevine f	666	5	10.9	357	1	ACTC_SOLTU	P30172	solanum tub
594	5	10.9	341	1	Y33E_MYCPN	P75302	mycoplasma	667	5	10.9	357	1	GLNA_PINSY	P52783	pinus sylve
595	5	10.9	341	1	YFTT_ECOLI	P39328	escherichia	668	5	10.9	357	1	LCP5_YEAST	P40079	saccharomyc
596	5	10.9	342	1	COBT_METSQ	Q9X7f4	methylobact	669	5	10.9	357	1	METX_HAEIN	P45131	haemophilus
597	5	10.9	342	1	ILVC_BACSU	P37253	bacillus su	670	5	10.9	357	1	YCIL_HAEIN	P45104	haemophilus
598	5	10.9	342	1	RFC3_SCHPO	O14003	schizosacch	671	5	10.9	357	1	YMR2_EBV	P03192	epstein-bar
599	5	10.9	342	1	Y755_METJA	Q58165	methanococc	672	5	10.9	358	1	ALF_BUCAI	P57526	buchnera ap
600	5	10.9	343	1	ALF_CORGL	P19537	corynebacte	673	5	10.9	358	1	ALF_ECOLI	P11604	escherichia
601	5	10.9	343	1	ALF_STRCO	Q9X8r6	streptomyce	674	5	10.9	358	1	ALF_EDWIC	Q52402	edwardsiell
602	5	10.9	343	1	BBUR_BORBR	O06703	bordelella	675	5	10.9	358	1	ALF_SCHPO	P36580	schizosacch
603	5	10.9	344	1	AR20_CABEL	Q18491	caenorhabdi	676	5	10.9	358	1	ALF_YEAST	P14540	saccharomyc
604	5	10.9	344	1	ARC2_TETH	O50146	thermus aqu	677	5	10.9	358	1	GLNA_LACSA	P23712	lactuca sat
605	5	10.9	344	1	Y613_METJA	Q58030	methanococc	678	5	10.9	359	1	ALF_BORBU	O51401	borrelia bu
606	5	10.9	344	1	YIBD_ECOLI	P11290	escherichia	679	5	10.9	359	1	ALF_HAEIN	P44429	haemophilus
607	5	10.9	345	1	ALF_MYCLE	O69600	mycobacteri	680	5	10.9	359	1	GPDA_DROKA	O97463	drosophila
608	5	10.9	345	1	EFB1_RAT	P25796	rattus norv	681	5	10.9	359	1	OP21_HAEIN	P43839	haemophilus
609	5	10.9	346	1	YIB9_YEAST	P40546	saccharomyc	682	5	10.9	359	1	RS2_SPICI	P19679	spiroplasma
610	5	10.9	347	1	FOS_CYPCA	P79702	cyprinus ca	683	5	10.9	360	1	ALF_DROME	P07764	drosophila
611	5	10.9	347	1	HPT_MUSSA	Q62558	mus saxicol	684	5	10.9	360	1	DCAM_SOLTU	Q04694	solanum tub
612	5	10.9	347	1	HPT_RAT	P08666	rattus norv	685	5	10.9	360	1	OP26_HAEIN	Q48216	haemophilus
613	5	10.9	347	1	Y4KE_RHISN	P55525	rhizobium s	686	5	10.9	361	1	DCAM_NICSY	O80402	nicotiana s
614	5	10.9	348	1	AROH_BUCAI	P57224	buchnera ap	687	5	10.9	361	1	DCAM_TOBAC	O40009	nicotiana t
615	5	10.9	348	1	AROH_BUCAP	P46245	buchnera ap	688	5	10.9	361	1	ELYA_BACHD	P41363	bacillus ha
616	5	10.9	348	1	AROH_ERWHE	O54459	erwinia her	689	5	10.9	361	1	OP22_HAEIN	P20149	haemophilus
617	5	10.9	348	1	PLGI_CAMJE	Q9pmj8	campylobact	690	5	10.9	361	1	OP29_HAEIN	Q48217	haemophilus

691	5	10.9	362	1	DCAM_DATST	Q96555	datastra	764	5	10.9	375	1	ACTC_STYPL	Q00215	styla plic
692	5	10.9	362	1	NSDL_MOUSE	P13100	mus musculus	765	5	10.9	375	1	ACTG_CEPAC	Q9uvw9	cephalospor
693	5	10.9	363	1	GPDA_DROME	P13706	drosophila	766	5	10.9	375	1	ACTG_EMENI	P20359	emericeella
694	5	10.9	363	1	OP2B_HAEN	Q48220	haemophilus	767	5	10.9	375	1	ACTG_HUMAN	P02571	homo sapien
695	5	10.9	363	1	ORC2_ARATH	Q38899	arabidopsis	768	5	10.9	375	1	ACTG_PENCH	Q9urs0	penicillium
696	5	10.9	364	1	AT9G_SP10L	P05435	spinacia ol	769	5	10.9	375	1	ACT_AJECA	P53455	ajelomyces
697	5	10.9	364	1	FNAL_BORBU	P28616	borrelia bu	770	5	10.9	375	1	ACT_BORCI	O13419	borryctis ci
698	5	10.9	364	1	DCTL_HUMAN	Q96a29	homo sapien	771	5	10.9	375	1	ACT_COPCI	Q9uvx4	coprinus ci
699	5	10.9	364	1	YB5_YEAST	P38241	saccharomyc	772	5	10.9	375	1	ACT_CRYNE	P48465	cryptococcu
700	5	10.9	365	1	RF2_ECOLI	P07012	escherichia	773	5	10.9	375	1	ACT_FUCDI	P53502	fucus disti
701	5	10.9	365	1	RF2_HAEN	P43918	haemophilus	774	5	10.9	375	1	ACT_NEUCR	P78711	neurospora
702	5	10.9	365	1	RF2_SALTY	P38353	salmonella	775	5	10.9	375	1	ACT_PHARKH	P53689	phaffia rho
703	5	10.9	366	1	ACTD_PHYPO	P24263	physarum po	776	5	10.9	375	1	ACT_PUGCR	P50138	puccinia gr
704	5	10.9	366	1	ALF_NEUCR	P53444	neurospora	777	5	10.9	375	1	ACT_SCHPO	P10989	schizosacch
705	5	10.9	366	1	MRYA_RHIME	Q52952	rhizobium m	778	5	10.9	375	1	ACT_THELA	P10365	thermomyces
706	5	10.9	366	1	YNT0_YEAST	P53938	saccharomyc	779	5	10.9	375	1	DEK_HUMAN	P35659	homo sapien
707	5	10.9	367	1	OP2A_HAEN	Q48219	haemophilus	780	5	10.9	375	1	DNAJ_ECOLI	P08822	escherichia
708	5	10.9	367	1	TISD_MOUSE	P23949	mus musculus	781	5	10.9	375	1	PER_DROCP	P91507	drosophila
709	5	10.9	368	1	ISPG_BUCAL	P57374	buchnera ap	782	5	10.9	375	1	PER_DROSC	P91705	drosophila
710	5	10.9	369	1	METB_HAEN	P44502	haemophilus	783	5	10.9	376	1	ACTI_AEDAE	P49128	aedes aegypt
711	5	10.9	369	1	OP27_HAEN	Q48218	haemophilus	784	5	10.9	376	1	ACTI_ARTSX	P18600	artemia sp.
712	5	10.9	369	1	PROB_NEIMA	Q91uk7	neisseria m	785	5	10.9	376	1	ACTI_BOMMO	P07836	bombyx mori
713	5	10.9	369	1	PROB_NEIMA	Q91zq2	neisseria m	786	5	10.9	376	1	ACTI_CABEL	P10983	caenorhabdi
714	5	10.9	369	1	RF2_THEMA	Q9x1r5	thermotoga	787	5	10.9	376	1	ACTI_DROME	P10987	drosophila
715	5	10.9	370	1	TRML_BACSU	Q35020	bacillus su	788	5	10.9	376	1	ACTI_HELER	P53462	heliocidar
716	5	10.9	370	1	YXPR_BACSU	P54957	bacillus su	789	5	10.9	376	1	ACTI_LUMTE	P92185	lytechinus
717	5	10.9	371	1	ACT2_NAEFO	P27132	naegleria f	790	5	10.9	376	1	ACTI_LYPTI	P30162	onchocerca
718	5	10.9	371	1	ACT2_DIPDE	P53458	diphyllobo	791	5	10.9	376	1	ACTI_ONCVO	P30164	pisum sativ
719	5	10.9	371	1	OP25_HAEN	P46027	haemophilus	792	5	10.9	376	1	ACTI_PEA	P22131	phytophthor
720	5	10.9	371	1	ROA1_HUMAN	P09651	homo sapien	793	5	10.9	376	1	ACTI_PHYIN	P10988	plasmodium
721	5	10.9	371	1	TRMU_BACHD	Q9kdf2	bacillus ha	794	5	10.9	376	1	ACTI_PLAFA	P41112	podocoryne
722	5	10.9	371	1	YI40_SNY3	P73408	synechocyst	795	5	10.9	376	1	ACTI_PODCA	O18499	saccoglossu
723	5	10.9	372	1	ACT_LUMRU	P91754	lumbricus r	796	5	10.9	376	1	ACTI_SACKO	P53470	schistosoma
724	5	10.9	373	1	ACT_CHOCR	P53499	chondrus cr	797	5	10.9	376	1	ACTI_SCHMA	P10990	strongyloce
725	5	10.9	373	1	BIOF_HELPJ	Q9zln3	helicobacte	798	5	10.9	376	1	ACTI_STRFN	P18601	artemia sp.
726	5	10.9	373	1	BIOF_HELPY	Q25320	helicobacte	799	5	10.9	376	1	ACT2_ARTSX	P45885	bactrocera
727	5	10.9	373	1	DMTL_HUMAN	Q9y5r6	homo sapien	800	5	10.9	376	1	ACT2_BACDO	P07837	bombyx mori
728	5	10.9	373	1	DP3B_MYCPU	Q98rk6	mycoplasma	801	5	10.9	376	1	ACT2_BOMMO	P10984	caenorhabdi
729	5	10.9	373	1	NSDL_HUMAN	Q15738	homo sapien	802	5	10.9	376	1	ACT2_CABEL	P53456	diphyllobo
730	5	10.9	373	1	RO31_XENLA	P51968	xenopus lae	803	5	10.9	376	1	ACT2_DIPDE	P2572	drosophila
731	5	10.9	373	1	US74_DROME	Q95y15	drosophila	804	5	10.9	376	1	ACT2_DROME	Q03341	echinococcu
732	5	10.9	374	1	ACTM_STRPU	P12431	strongyloce	805	5	10.9	376	1	ACT2_ECHGR	P02176	lumbricus t
733	5	10.9	374	1	YF18_METJA	Q58913	methanococc	806	5	10.9	376	1	ACT2_LUMTE	P53466	lytechinus
734	5	10.9	375	1	ACT1_ACACA	P02578	acanthamoeb	807	5	10.9	376	1	ACT2_LYPTI	P30163	onchocerca
735	5	10.9	375	1	ACT1_DICDI	P02577	dictyosteli	808	5	10.9	376	1	ACT2_ONCVO	P30165	pisum sativ
736	5	10.9	375	1	ACT1_ECHGR	P53432	echinococcu	809	5	10.9	376	1	ACT2_PEA	P14883	plasmodium
737	5	10.9	375	1	ACT1_FUGRU	P53484	fugu rubrip	810	5	10.9	376	1	ACT2_PLAFA	O18500	saccoglossu
738	5	10.9	375	1	ACT1_MAIZE	P02582	zea mays (m	811	5	10.9	376	1	ACT2_SACKO	P53471	schistosoma
739	5	10.9	375	1	ACT1_NAEFO	P27131	naegleria f	812	5	10.9	376	1	ACT2_SCHMA	P10993	tetrahymena
740	5	10.9	375	1	ACT1_OXYNO	P12715	oxytricha n	813	5	10.9	376	1	ACT2_TETPY	P45886	bactrocera
741	5	10.9	375	1	ACT1_OXYTR	P53468	oxytricha t	814	5	10.9	376	1	ACT3_BACDO	P04829	bombyx mori
742	5	10.9	375	1	ACT1_PNECA	P43239	pneumocysti	815	5	10.9	376	1	ACT3_BOMMO	P53501	helicoverpa
743	5	10.9	375	1	ACT1_SCHCO	Q9y702	schizophyll	816	5	10.9	376	1	ACT3_DROME	P41340	limulus pol
744	5	10.9	375	1	ACT1_SUIBO	Q9y701	suillus bov	817	5	10.9	376	1	ACT3_HELAM	P17299	oryza sativ
745	5	10.9	375	1	ACT1_TETTH	P10992	tetrahymena	818	5	10.9	376	1	ACT3_LIMPO	P41113	podocoryne
746	5	10.9	375	1	ACT2_DICDI	P07827	dictyosteli	819	5	10.9	376	1	ACT3_ORYSA	P02580	glycine max
747	5	10.9	375	1	ACT2_FUGRU	P53485	fugu rubrip	820	5	10.9	376	1	ACT3_PODCA	P18603	artemia sp.
748	5	10.9	375	1	ACT2_OXYNO	P55805	oxytricha n	821	5	10.9	376	1	ACT3_SOYBN	Q27250	bombyx mori
749	5	10.9	375	1	ACT2_OXYTR	P53469	oxytricha t	822	5	10.9	376	1	ACT4_ARTSX	P10986	caenorhabdi
750	5	10.9	375	1	ACT2_SCHCO	Q9y896	schizophyll	823	5	10.9	376	1	ACT4_BOMMO	P20574	drosophila
751	5	10.9	375	1	ACT2_SUIBO	Q9y707	suillus bov	824	5	10.9	376	1	ACT4_CABEL	P53501	helicoverpa
752	5	10.9	375	1	ACT3_DICDI	P07829	dictyosteli	825	5	10.9	376	1	ACT4_DROME	P41340	limulus pol
753	5	10.9	375	1	ACT3_FUGRU	P53486	fugu rubrip	826	5	10.9	376	1	ACT5_BACDO	P17299	oryza sativ
754	5	10.9	375	1	ACT3_DICDI	P07830	dictyosteli	827	5	10.9	376	1	ACT5_CHICK	P41113	podocoryne
755	5	10.9	375	1	ACTA_PHYPO	P02576	physarum po	828	5	10.9	376	1	ACT5_DROME	P02580	glycine max
756	5	10.9	375	1	ACTB_CRIGR	P48975	cricetusul	829	5	10.9	376	1	ACT5_XENLA	P17300	oryza sativ
757	5	10.9	375	1	ACTB_GVPCA	P12714	cyprinus ca	830	5	10.9	376	1	ACT6_DROME	P53506	xenopus lae
758	5	10.9	375	1	ACTB_HUMAN	P02570	homo sapien	831	5	10.9	376	1	ACT7_ORYSA	P41339	limulus pol
759	5	10.9	375	1	ACTB_RABIT	P29751	oryctolagus	832	5	10.9	376	1	ACT8_XENLA	P53472	strongyloce
760	5	10.9	375	1	ACTB_SALSA	Q42161	salmo salar	833	5	10.9	376	1	ACTA_LIMPO	P53473	strongyloce
761	5	10.9	375	1	ACTB_BRABE	Q93129	branchiosto	834	5	10.9	376	1	ACTA_STRPU	P53461	halocynthia
762	5	10.9	375	1	ACTC_BRABE	Q93131	branchiosto	835	5	10.9	376	1	ACTB_STRPU		
763	5	10.9	375	1	ACTC_BRALA	O17503	branchiosto	836	5	10.9	376	1	ACTC_HAIRO		

837	5	10.9	376	1	ACTC_PISOC	P12716	pisaster oc	910	5	10.9	378	1	ACTM_BRAFL	Q93132	branchiost
838	5	10.9	376	1	ACTC_STRPU	Q07903	strongyloce	911	5	10.9	378	1	ACTM_CIOSA	O15998	ciona savi9
839	5	10.9	376	1	ACTD_STRPU	P10991	strongyloce	912	5	10.9	378	1	ACTM_MOLOC	P53467	molgula ocu
840	5	10.9	376	1	ACTE_STRPU	P53474	strongyloce	913	5	10.9	378	1	ACTM_STYCL	P26198	styla clav
841	5	10.9	376	1	ACTF_STRPU	P18499	strongyloce	914	5	10.9	378	1	ACTN_STYCL	P53475	styla clav
842	5	10.9	376	1	ACTH_HUMAN	P12718	homo sapien	915	5	10.9	378	1	ACT_SCHDU	O65314	scherrfella
843	5	10.9	376	1	ACTM_APLCA	P17304	alyxia cal	916	5	10.9	378	1	DNAJ_SALTY	Q60004	salmonella
844	5	10.9	376	1	ACTM_HELER	P53463	helioaidari	917	5	10.9	378	1	DP3B_STRPN	O66672	streptococc
845	5	10.9	376	1	ACTM_PISOC	P53464	helioaidari	918	5	10.9	378	1	O33A_DROME	P15144	drosofila
846	5	10.9	376	1	ACTM_PISOC	P12717	pisaster oc	919	5	10.9	378	1	YDCC_ECOLI	P28917	escherichia
847	5	10.9	376	1	ACTC_FUGRU	P53483	fugu rubrip	920	5	10.9	378	1	YH1I_ECOLI	P28912	escherichia
848	5	10.9	376	1	ACTX_LIMPO	P14341	limulus pol	921	5	10.9	378	1	YH1I_ECOLI	P46080	anabaena sp
849	5	10.9	376	1	ACT_ACHBI	P26182	achlya bise	922	5	10.9	379	1	YH1I_ECOLI	P17298	oryza sativ
850	5	10.9	376	1	ACT_BIOGL	P92179	biomphalari	923	5	10.9	379	1	ACTM_BRABE	Q93130	branchiost
851	5	10.9	376	1	ACT_BRUMA	P90689	brugia mala	924	5	10.9	379	1	ACTM_STYPL	Q00214	styla plic
852	5	10.9	376	1	ACT_CANAL	P14235	candida alb	925	5	10.9	379	1	ACTM_STYPL	O50161	borrelia bu
853	5	10.9	376	1	ACT_CANDU	O94928	candida dub	926	5	10.9	380	1	Y374_BORBU	P23343	daucus caro
854	5	10.9	376	1	ACT_CRAIG	O17320	crassostrea	927	5	10.9	380	1	ACT1_DAUCA	P23344	daucus caro
855	5	10.9	376	1	ACT_CRYPV	P26183	cryptospori	928	5	10.9	380	1	ACTM_BRALA	O17502	branchiost
856	5	10.9	376	1	ACT_ENTHI	P11426	entamoeba h	929	5	10.9	380	1	DNAJ_BACST	Q45552	bacillus st
857	5	10.9	376	1	ACT_FUCVE	Q39758	fucus vesic	930	5	10.9	380	1	ME1B_HELPJ	Q92947	helicobacte
858	5	10.9	376	1	ACT_HYDAT	P17126	hydra atten	931	5	10.9	380	1	ME1B_HELPJ	P56069	helicobacte
859	5	10.9	376	1	ACT_MANSE	P49871	manduca sex	932	5	10.9	380	1	PLA_COLGL	Q00374	colletotric
860	5	10.9	376	1	ACT_MAYDE	O16808	mayetiola d	933	5	10.9	381	1	POLG_MDMV	P32652	maize dwarf
861	5	10.9	376	1	ACT_PICAN	Q14258	pichia anqu	934	5	10.9	381	1	ACT2_DAUCA	P23344	daucus caro
862	5	10.9	376	1	ACT_PLANG	Q26065	plascopeten	935	5	10.9	382	1	COAL_POVNA	P03090	mouse polyo
863	5	10.9	376	1	ACT_TAESO	P14227	taenia soli	936	5	10.9	382	1	RS1H_BACSU	P38494	bacillus su
864	5	10.9	376	1	ACT_TOXGO	P53476	toxoplasma	937	5	10.9	383	1	COAL_POVNA	P03091	mouse polyo
865	5	10.9	376	1	RE5C_HUMAN	Q43502	homo sapien	938	5	10.9	383	1	COAL_POVNC	P12907	mouse polyo
866	5	10.9	376	1	RE5C_HUMAN	P42442	corynebacte	939	5	10.9	383	1	COAL_POVNC	P42902	mouse polyo
867	5	10.9	376	1	XYNA_BACOV	P49942	bacteroides	940	5	10.9	385	1	ACT2_PNECA	P42023	pneumocysti
868	5	10.9	376	1	ACT1_ORYLA	O98972	oryzias lat	941	5	10.9	385	1	OMP_BORPE	Q04064	bordeletia
869	5	10.9	377	1	ACT1_ORYSA	P13362	oryza sativ	942	5	10.9	385	1	OP23_HAEIN	P46025	haemophilus
870	5	10.9	377	1	ACT1_SORBI	P53504	sorghum bic	943	5	10.9	385	1	PER_DRONE	P16866	drosofila
871	5	10.9	377	1	ACT1_SOYBN	P02581	glycine max	944	5	10.9	385	1	Y464_MYCCE	P47702	mycoplasma
872	5	10.9	377	1	ACT1_TOBAC	Q05214	nicotiana t	945	5	10.9	386	1	OP24_HAEIN	P46026	haemophilus
873	5	10.9	377	1	ACT2_XENIA	P44751	xenopus lae	946	5	10.9	387	1	TIZ2_MOUSE	Q9eqn3	mus musculus
874	5	10.9	377	1	ACT2_XENIA	P26197	absidia gla	947	5	10.9	388	1	GALL_STACA	O9rgs1	staphylococ
875	5	10.9	377	1	ACT2_XENIA	P53492	arabidopsis	948	5	10.9	388	1	PERC_CALJA	O9n3d3	callithrix
876	5	10.9	377	1	ACT2_XENIA	P10995	xenopus lae	949	5	10.9	388	1	RECA_STRPN	P20142	homo sapien
877	5	10.9	377	1	ACT2_XENIA	P20399	xenopus tro	950	5	10.9	388	1	PERC_STRPN	P30758	streptococc
878	5	10.9	377	1	ACT2_XENIA	P53493	arabidopsis	951	5	10.9	389	1	BIOT_BACSH	P22806	bacillus sp
879	5	10.9	377	1	ACT3_DIPDE	P53457	diphylobot	952	5	10.9	389	1	DNAJ_METWA	P25815	methanosarc
880	5	10.9	377	1	ACT3_PEA	P46258	pisum sativ	953	5	10.9	390	1	DFF_BORBU	O51752	borrelia bu
881	5	10.9	377	1	ACT3_SOLTU	P30167	solanum tub	954	5	10.9	390	1	PER_DROTP	P23256	escherichia
882	5	10.9	377	1	ACT3_XENIA	P04752	xenopus lae	955	5	10.9	390	1	PER_DROTP	P91716	drosofila
883	5	10.9	377	1	ACT4_ARATH	P53494	arabidopsis	956	5	10.9	390	1	Y4RI_RHISN	P19717	mumps virus
884	5	10.9	377	1	ACT6_SOLTU	P30168	solanum tub	957	5	10.9	391	1	CATE_CAVPO	P55642	rhizobium s
885	5	10.9	377	1	ACT7_SOLTU	P30169	solanum tub	958	5	10.9	391	1	EFTU_RHILU	P25796	cavia porce
886	5	10.9	377	1	ACTA_CHICK	P08023	gallus gall	959	5	10.9	391	1	GALL_MOUSE	O981f7	rhizobium l
887	5	10.9	377	1	ACTA_HUMAN	P03996	homo sapien	960	5	10.9	391	1	PER_DRONIN	O90n0	mus musculus
888	5	10.9	377	1	ACTB_ARATH	P53496	arabidopsis	961	5	10.9	391	1	RRPP_MUMPE	P91613	drosofila
889	5	10.9	377	1	ACTB_ARATH	P30171	solanum tub	962	5	10.9	391	1	RRPP_MUMPE	P16072	mumps virus
890	5	10.9	377	1	ACTC_ARATH	P53497	arabidopsis	963	5	10.9	392	1	PAF2_BOVIN	P16595	mumps virus
891	5	10.9	377	1	ACTC_FUGRU	P53480	fugu rubrip	964	5	10.9	392	1	PAF2_HUMAN	P79106	bos taurus
892	5	10.9	377	1	ACTC_HUMAN	P04270	homo sapien	965	5	10.9	392	1	SB11_HUMAN	O99487	homo sapien
893	5	10.9	377	1	ACTD_SOLTU	P30173	solanum tub	966	5	10.9	392	1	TRPB_BUCSC	Q96pl5	homo sapien
894	5	10.9	377	1	ACTS_CARAU	P49055	carassius a	967	5	10.9	392	1	Y462_TREPA	O59169	buchnera ap
895	5	10.9	377	1	ACTS_CYPCA	P53479	cyprinus ca	968	5	10.9	393	1	ACKA_MYCCE	O83475	treponema p
896	5	10.9	377	1	ACTS_FUGRU	P53481	fugu rubrip	969	5	10.9	393	1	GCIM_MOUSE	P47599	mycoplasma
897	5	10.9	377	1	ACTS_HUMAN	P02568	homo sapien	970	5	10.9	393	1	MT04_HELPY	P01869	mus musculus
898	5	10.9	377	1	ACTT_FUGRU	P53482	fugu rubrip	971	5	10.9	393	1	PGK_BORBU	O25443	helicobacte
899	5	10.9	377	1	ACTT_CHLRE	P53498	chlamydomon	972	5	10.9	394	1	CIM3_HUMAN	Q14649	homo sapien
900	5	10.9	377	1	ACTT_COLSC	O65315	coleochaete	973	5	10.9	394	1	YHFF_BACSU	Q13115	homo sapien
901	5	10.9	377	1	ACT_CYAME	P53500	cyanidiosch	974	5	10.9	394	1	NEUA_STRAG	P39642	bacillus su
902	5	10.9	377	1	ACT_MESVI	O65316	mesostigma	975	5	10.9	395	1	PER_DRONE	Q53598	streptococc
903	5	10.9	377	1	ACT_VOLCA	P20904	volvox cart	976	5	10.9	395	1	PER_DRONE	P41999	caenorhabdi
904	5	10.9	377	1	DNAJ_BUCAL	O32465	buchnera ap	977	5	10.9	395	1	TI22_HUMAN	P92203	drosofila
905	5	10.9	377	1	OR5B_APLINE	P90680	apis mellif	978	5	10.9	395	1	Y414_METUA	O9y3a8	homo sapien
906	5	10.9	377	1	PERC_MACRU	P03955	macaca fusc	979	5	10.9	395	1	Y414_METUA	O57857	methanococc
907	5	10.9	378	1	ACT1_HALRO	P53460	halocynthia	980	5	10.9	396	1	APR4_HUMAN	P06727	homo sapien
908	5	10.9	378	1	ACT2_HALRO	P27130	halocynthia	981	5	10.9	396	1	BRB2_RAT	P25023	rattus norv
909	5	10.9	378	1	ACT2_MOLOC	Q25472	molgula ocu	982	5	10.9	396	1	CATE_RABIT	P43159	oryctolagus

```

983 5 10.9 396 1 PER_DROPV P01697 drosophila
984 5 10.9 396 1 PER_DROPV P01698 drosophila
985 5 10.9 396 1 PER_DROPV Q2554 staphylococ
986 5 10.9 396 1 PER_DROPV P4535 haemophilus
987 5 10.9 396 1 SOTB_HAEN Q12185 saccharomyc
988 5 10.9 396 1 YWB_BACSU P39587 bacillus su
989 5 10.9 396 1 6P22_YEAST Q12471 saccharomyc
990 5 10.9 397 1 CATE_MOUSE P70269 mus musculus
991 5 10.9 397 1 CATE_MOUSE Q08339 erythrocebu
992 5 10.9 397 1 DJM_MOUSE Q03mc3 mus musculus
993 5 10.9 397 1 LHX3_HUMAN Q9ubf4 homo sapien
994 5 10.9 397 1 THIM_HUMAN P42765 homo sapien
995 5 10.9 397 1 THIM_RAT P13437 rattus norv
996 5 10.9 398 1 APL1_HUMAN O14791 homo sapien
997 5 10.9 398 1 CATE_RAT P16228 rattus norv
998 5 10.9 398 1 DAP3_HUMAN P51398 homo sapien
999 5 10.9 398 1 TRA5_RHIME Q52873 rhizobium m
1000 5 10.9 399 1 GCAM_MOUSE P01865 mus musculus

```

ALIGNMENTS

```

RESULT 1
Y228_BORBU STANDARD; PRT; 971 AA.
AC O51246;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein BB0228.
GN BB0228.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uterback T., Wattley L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi."
RL Nature 390:580-586(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE001133; AAC66621.1; -
CC TIGR; BB0228; -
CC KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 971 AA; 112959 MW; 08BA688D7B8C591A CRC64;

```

```

Query Match 19.6%; Score 9; DB 1; Length 971;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 TSLKNDKF 12
| | | | | | |
Db 653 TSLKNDKF 661

```

```

RESULT 2
FILL_YEAST STANDARD; PRT; 230 AA.
ID FILL_YEAST
AC P38771;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE FILL protein, mitochondrial precursor.
GN FILL OR KIM4 OR YHR038W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MT8-1;
RX MEDLINE=984117448; PubMed=9746366;
RA Kanai T., Takeshita S., Atomi H., Umemura K., Ueda M., Tanaka A.;
RT "A regulatory factor, FILL, involved in derepression of the
RT isocitrate lyase gene in Saccharomyces cerevisiae -- a possible
RT mitochondrial protein necessary for protein synthesis in
RT mitochondria."
RL Eur. J. Biochem. 256:212-220(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba P., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nhan M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."
RL Science 265:2077-2082(1994).
RN [3]
RP SIMILARITY TO RRF.
RX MEDLINE=96155066; PubMed=8563640;
RA Ouzounis C., Bork P., Casari G., Sander C.;
RT "New protein functions in yeast chromosome VIII."
RL Protein Sci. 4:2424-2428(1995).
CC -1- FUNCTION: INVOLVED IN GLUCOSE REPRESSION/DEREPRESSION. MAY BE
CC NECESSARY FOR PROTEIN SYNTHESIS IN MITOCHONDRIA. MAY FUNCTION AS A
CC RIBOSOME RECYCLING FACTOR IN MITOCHONDRIA.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC -1- SIMILARITY: BELONGS TO THE RRF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB016033; BAA31687.1; -
CC EMBL; J00062; AAB68906.1; -
CC PIR; S46737; S46737.
CC SGD; S0001080; FILL.
CC InterPro; IPR002661; RRF.
CC Pfam; PF01765; RRF.
CC KW Protein biosynthesis; Transist peptide; Mitochondrion.
FT TRANSIT 1 24 MITOCHONDRION (POTENTIAL).
FT CHAIN 25 230 FILL PROTEIN.
SQ SEQUENCE 230 AA; 26406 MW; 9CB8C85F86F3008 CRC64;

```

```

Query Match 15.2%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 34 RLFRNSF 40
```



```

Db      17 RLFRNSF 23
|||||||
RESULT 3
ABFL_TRIKE
ID ABFL_TRIKE STANDARD; PRT; 500 AA.
AC Q92455;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-L-arabinofuranosidase precursor (EC 3.2.1.55) (Arabinosidase).
GN ABFL.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QM9414 / RUT C-30;
RX MEDLINE=96434473; PubMed=8837440;
RA Margolles-Clark E., Tenkanen M., Nakari-Setäläe T., Penttiläe M.;
RT "Cloning of genes encoding alpha-L-arabinofuranosidase and beta-
RT xylosidase from Trichoderma reesei by expression in Saccharomyces
RT cerevisiae.";
RL Appl. Environ. Microbiol. 62:3840-3846(1996).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing alpha-L-
CC arabinofuranoside residues in alpha-L-arabinosides.
CC -!- PATHWAY: INVOLVED IN DEGRADATION OF THE PLANT CELL WALL
CC POLYSACCHARIDE L-ARABINAN.
CC -!- SIMILARITY: BELONGS TO FAMILY 54 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z69252; CAA93243.1; -.
KW Hydrolyase; Glycosidase; Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 500
FT CARBOHYD 467 467
FT SEQUENCE 500 AA; 5115 MW; FC6B4DB03EE3C762 CRC64;
SQ SEQUENCE 500 AA; 5115 MW; FC6B4DB03EE3C762 CRC64;

Query Match 15.2%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 SGLSTF 32
Db 266 SGLSTF 272
|||||||
RESULT 4
XYLL_TRIKO
ID XYLL_TRIKO STANDARD; PRT; 500 AA.
AC P48792;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arabinofuranosidase/B-xylosidase precursor (Includes: Alpha-L-
DE arabinofuranosidase (EC 3.2.1.55) (Arabinosidase); Beta-xylosidase
DE (EC 3.2.1.37) (1,4-beta-D-xylan xylohydrolase) (Xylan 1,4-beta-
DE xylosidase)).
GN XYLL.
OS Trichoderma koningii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=55202;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=G-39;
RA Huang L.N., Hseu T.H., Lee Y.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing alpha-L-
CC arabinofuranoside residues in alpha-L-arabinosides.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-XYLANS SO AS TO
CC REMOVE SUCCESSIVE D-XYLOSE RESIDUES FROM THE NON-REDUCING
CC TERMINI. IT ALSO HYDROLYSES XYLOBIOSE.
CC -!- SIMILARITY: BELONGS TO FAMILY 54 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38661; AAA81024.1; -.
KW Xylan degradation; Hydrolase; Glycosidase; Multifunctional enzyme;
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 500
FT CARBOHYD 467 467
FT SEQUENCE 500 AA; 51129 MW; 06DFC319AFA1149 CRC64;
SQ SEQUENCE 500 AA; 51129 MW; 06DFC319AFA1149 CRC64;

Query Match 15.2%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 SGLSTF 32
Db 266 SGLSTF 272
|||||||
RESULT 5
RELB_MOUSE
ID RELB_MOUSE STANDARD; PRT; 558 AA;
AC Q04863;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription factor RELB.
GN RELB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92123192; PubMed=1732739;
RA Ryseck R.P., Bull P., Takamiya M., Bours V., Siebenlist U.,
RA Dobrzanski P., Bravo R.;
RT "RelB, a new Rel family transcription activator that can interact
RT with p50-NF-kappa B.";
RL Mol. Cell. Biol. 12:674-684(1992).
RN [2]
RP SEQUENCE OF 1-116 FROM N.A.
RX MEDLINE=93180804; PubMed=8441398;
RA Dobrzanski P., Ryseck R.P., Bravo R.;
RT "Both N- and C-terminal domains of RelB are required for full
RT transactivation: role of the N-terminal leucine zipper-like motif.";
RL Mol. Cell. Biol. 13:1572-1582(1993).
RN [3]
RP SEQUENCE OF 309-429 FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Liver;
RX MEDLINE=95147976; PubMed=7845467;
RA Burkly L., Hession C., Ogata L., Reilly C., Marconi L.A.,
RA Olson D., Tizard R., Cate R., Io D.;
RT "Expression of relB is required for the development of thymic medulla
RT and dendritic cells.";
RL Nature 373:531-536(1995).

```

CC -!- FUNCTION: STIMULATES PROMOTER ACTIVITY IN THE PRESENCE OF P49- AND
CC P50-NF-KAPPA-B. NEITHER ASSOCIATES WITH DNA NOR WITH P65-NF-KAPPA-
CC B.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN INTESTINE, THYMUS AND SPLEEN.
CC UNDETECTABLE IN LIVER, BONE MARROW, KIDNEY AND TESTIS.
CC -!- DOMAIN: BOTH N- AND C-TERMINAL DOMAINS ARE REQUIRED FOR
CC TRANSCRIPTIONAL ACTIVATION.
CC -!- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83380; AAA40041.1; -;
CC EMBL; S56076; AAB35493.2; -;
CC EMBL; S76754; AAB33259.1; -;
CC PIR; A42023; A42023.
CC HSP; P25799; 1BFT.
CC TRANSFAC; T01932; -;
CC MGI; 103289; Relb.
CC InterPro; IPR002909; IPT_TIG.
CC InterPro; IPR000451; REL.
CC Pfam; PF00554; RHD; 1.
CC Pfam; PF01833; TIG; 1.
CC PRINTS; PR00057; NFKB1NSCPFT.
CC SMART; SM00429; IPT; 1.
CC PROSITE; PS01204; REL_1; 1.
CC PROSITE; PS0254; REL_2; 1.
CC Nuclear protein; Transcription regulation; Activator; Phosphorylation.
CC DOMAIN 22 50
CC REL-LIKE (RHD).
CC FT DOMAIN 103 418
CC FT DOMAIN 387 391
CC FT DOMAIN 411 416
CC FT CONFLICT 51 51
CC FT V -> D (IN REF. 2).
CC SQ SEQUENCE 558 AA; 60304 MW; 1ED2A354C6EDD3D CRC64;

CC Query Match 15.2%; Score 7; DB 1; Length 558;
CC Best Local Similarity 100.0%; Pred. No. 11;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC QY 16 TDELEII 22
CC DB 33 TDELEII 39

CC RESULT 6
CC RELB_HUMAN STANDARD; PRT; 579 AA.
CC AC Q01201;
CC DT 01-OCT-1993 (Rel. 27, Created)
CC DT 01-OCT-1993 (Rel. 27, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Transcription factor RELB (1-Rel).
CC GN RELB.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE-T-cell;
CC RX MEDLINE=92249768; PubMed=1577270;
CC RA Ruben S.M., Klement J.F., Maher M., Coleman T.A., Chen C.H.,
CC RA Rosen C.A.;
CC FT "1-Rel: a novel rel-related protein that inhibits NF-kappa B
CC FT transcriptional activity."
CC RL Genes Dev. 6:745-760(1992).
CC RN [2]

RP FUNCTION.
RX MEDLINE=93180804; PubMed=8441398;
RA Dobrzanski P., Kyssek R.P., Bravo R.;
RT "Both N- and C-terminal domains of Relb are required for full
RT transactivation: role of the N-terminal leucine zipper-like motif";
RL Mol. Cell. Biol. 13:1572-1582(1993).
CC -!- FUNCTION: STIMULATES PROMOTER ACTIVITY IN THE PRESENCE OF P49- AND
CC P50-NF-KAPPA-B. NEITHER ASSOCIATES WITH DNA NOR WITH P65-NF-KAPPA-
CC B.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- INDUCTION: BY MITOGENS.
CC -!- DOMAIN: BOTH N- AND C-TERMINAL DOMAINS ARE REQUIRED FOR
CC TRANSCRIPTIONAL ACTIVATION.
CC -!- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO INHIBIT THE
CC TRANSCRIPTIONAL ACTIVITY OF NUCLEAR FACTOR NF-KAPPA-B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M83221; AAA36127.1; -;
CC PIR; A42617; A42617.
CC HSP; P25799; 1BFT.
CC TRANSFAC; T01931; -;
CC MIM; 604758; -;
CC InterPro; IPR002909; IPT_TIG.
CC InterPro; IPR000451; REL.
CC Pfam; PF00554; RHD; 1.
CC Pfam; PF01833; TIG; 1.
CC PRINTS; PR00057; NFKB1NSCPFT.
CC SMART; SM00429; IPT; 1.
CC PROSITE; PS01204; REL_1; 1.
CC PROSITE; PS0254; REL_2; 1.
CC Nuclear protein; Transcription regulation; Activator; Phosphorylation.
CC DOMAIN 40 68
CC REL-LIKE (RHD).
CC FT DOMAIN 125 440
CC FT DOMAIN 433 438
CC FT NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SQ SEQUENCE 579 AA; 61990 MW; 5BFFBF8A54184D52 CRC64;

CC Query Match 15.2%; Score 7; DB 1; Length 579;
CC Best Local Similarity 100.0%; Pred. No. 11;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC QY 16 TDELEII 22
CC DB 51 TDELEII 57

CC RESULT 7
CC YN48 YEAST STANDARD; PRT; 591 AA.
CC AC P42846;
CC DT 01-NOV-1995 (Rel. 32, Created)
CC DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC DT 01-OCT-1996 (Rel. 34, Last annotation update)
CC DE Hypothetical 68.7 kDa protein in S7B1-MCK1 intergenic region.
CC GN YNL308C OR N0388.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
CC OX NCBI_TaxID=4932;
CC [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=S288C / FY1676;
CC RX MEDLINE=96076632; PubMed=7502583;
CC RA Maftahi M., Nicaud J.-M., Levesque H., Gaillardin C.;
CC RT "Sequencing analysis of a 24.7 kb fragment of yeast chromosome XIV
CC RT identifies six known genes, a new member of the hexose transporter

RT family and ten new open reading frames.";
 CC Yeast 11:1077-1085(1995).
 CC -!- SIMILARITY: SOME, TO S.POMBE SPAC22G7.05.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 246259; CAA8387.1; -
 CC DR EMBL; 271584; CAA96237.1; -
 CC DR SGD; S000522; YNL308C.
 CC KW Hypothetical protein.
 CC FT DOMAIN 53 65 POLY-GLU.
 CC FT DOMAIN 457 465 POLY-GLU.
 CC SQ SEQUENCE 591 AA; 68653 MW; 38637571EDB05EA6 CRC64;
 CC -----
 CC Query Match 15.2%; Score 7; DB 1; Length 591;
 CC Best Local Similarity 100.0%; Pred. No. 12;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 7 LKNDKE 13
 CC DB 352 LKNDKE 358
 CC -----
 CC RESULT 8
 CC UGS3_SOLITU STANDARD; PRT; 788 AA.
 CC ID UGS3_SOLITU
 CC AC Q43847;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 01-MAR-2002 (Rel. 41, Last annotation update)
 CC DE Glycogen (starch) synthase, chloroplast precursor (EC 2.4.1.11)
 CC DE (GBSSI) (Granule-bound starch synthase II) (Fragment).
 CC OS Solanum tuberosum (Potato).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 CC OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 CC OX NCBI_TaxID=4113;
 CC [1]
 CC SEQUENCE FROM N.A., AND SEQUENCE OF 66-75.
 CC RP STRAIN=CV. DESIREE; TISSUE=tuber;
 CC RX MEDLINE=95400340; PubMed=7670507;
 CC RA Edwards A., Marshall J., Sidebottom C., Visser R.G.F., Smith A.M.,
 CC Martin C.;
 CC RT "Biochemical and molecular characterization of a novel starch
 CC synthase from potato tubers.";
 CC RL Plant J. 8:283-294(1995).
 CC -!- FUNCTION: ACCOUNTS FOR ONLY 10 TO 15% OF THE TOTAL SOLUBLE STARCH
 CC -!- SYNTHASE ACTIVITY IN TUBERS.
 CC CC CATALYTIC ACTIVITY: UDP-glucose + [(1,4)-alpha-D-glucosyl](N) =
 CC UDP + [(1,4)-alpha-D-glucosyl](N+1).
 CC -!- PATHWAY: STARCH BIOSYNTHESIS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST/AMYOPLAST, SOLUBLE AND GRANULE-
 CC BOUND.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X87988; CAA61241.1; -
 CC DR InterPro; IPR001296; Glycos_transf_1.
 CC Pfam; PF00534; Glycos_transf_1; 1.

KW Glycogen biosynthesis; Transferase; Glycosyltransferase;
 KW Transit peptide; Chloroplast; Amyloplast; Starch biosynthesis.
 FT NON_TER 1
 FT TRANSIT <1 65 CHLOROPLAST.
 FT CHAIN 66 788 GLYCOGEN [STARCH] SYNTHASE.
 FT BINDING 310 310 UDP-GLUCOSE (BY SIMILARITY).
 FT VARIANT 71 71 S -> D.
 SQ SEQUENCE 788 AA; 87890 MW; 8D8B90611E862B7B CRC64;
 CC -----
 CC Query Match 15.2%; Score 7; DB 1; Length 788;
 CC Best Local Similarity 100.0%; Pred. No. 15;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 24 GSGSLS 30
 CC DB 128 GSGSLS 134
 CC -----
 CC RESULT 9
 CC K6PP_RABIT STANDARD; PRT; 791 AA.
 CC ID K6PP_RABIT
 CC AC P47859;
 CC DT 01-FEB-1996 (Rel. 33, Created)
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE 6-phosphofructokinase, type C (EC 2.7.1.11) (Phosphofructokinase
 CC DE 1) (Phosphohexokinase) (Phosphofructo-1-kinase isozyme C) (PFK-C).
 CC GN PFKP.
 CC OS Oryctolagus cuniculus (Rabbit).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC OX NCBI_TaxID=9986;
 CC [1]
 CC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 CC RP TISSUE=Brain;
 CC RX MEDLINE=94164929; PubMed=81119919;
 CC RA Li Y., Valaitis A.P., Latshaw S.P., Kwiatkowska D., Tripathi R.L.,
 CC Campbell M.C., Kemp R.G.;
 CC RT "Structure and expression of the cDNA for the C isozyme of
 CC phosphofructo-1-kinase from rabbit brain.";
 CC RL J. Biol. Chem. 269:5781-5787(1994).
 CC CC CATALYTIC ACTIVITY: ATP + D-fructose 6-phosphate = ADP + D-
 CC fructose 1,6-bisphosphate.
 CC CC COFACTOR: MAGNESIUM.
 CC CC ENZYME REGULATION: ALLOSTERIC ENZYME ACTIVATED BY ADP, AMP, OR
 CC FRUCTOSE BISPHOSPHATE AND INHIBITED BY ATP OR CITRATE.
 CC -!- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY. TWO DOMAIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U01154; AAA17707.1; -
 CC DR HSSP; P00512; 3PFK.
 CC DR InterPro; IPR000023; Phosphofructokinase.
 CC Pfam; PF00365; PFK; 4.
 CC DR PRINTS; PR00476; PFRCTKINASE.
 CC DR PRODOM; PD000707; PHOSPHOFRUCTOKINASE; 2.
 CC DR PROSITE; PS00433; PHOSPHOFRUCTOKINASE; 2.
 CC KW Kinase; Transferase; Glycolysis; Repeat; Allosteric enzyme;
 CC KW Phosphorylation; Magnesium; Multigene family;
 CC SQ SEQUENCE 791 AA; 86349 MW; 3C10A36F229FD8E8 CRC64;
 CC -----
 CC Query Match 15.2%; Score 7; DB 1; Length 791;
 CC Best Local Similarity 100.0%; Pred. No. 15;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 IGGSGSL 29
Db 125 IGGSGSL 131

RESULT 10
PSD8_DICDI STANDARD; PRT; 88 AA.
AC P02889; P09409;
DT 21-JUL-1986 (Rel. 01, Created)
DT 16-OCT-1996 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Vegetative cell protein X (M4 protein) (Fragment).
GN M4.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81124286; PubMed=7465423;
RA Kimmel A.R., Firtel R.A.;
RT "Intervening sequences in a Dictyostellium gene that encodes a low
abundance class mRNA."
RL Nucleic Acids Res. 8:5599-5610(1980).
RN [2]
RP SEQUENCE OF 1-12 FROM N.A.
RX MEDLINE=86310885; PubMed=2427932;
RA Kimmel A.R., Firtel R.A.;
RT "Sequence organization and developmental expression of an
interspersed, repetitive element and associated single-copy DNA
sequences in Dictyostellium discoideum."
RL Mol. Cell. Biol. 5:2123-2130(1985).
CC -1- FUNCTION: ACTS AS A REGULATORY SUBUNIT OF THE 26 PROTEASOME WHICH
IS INVOLVED IN THE ATP-DEPENDENT DEGRADATION OF UBIQUITINATED
PROTEINS (BY SIMILARITY).
CC -1- DEVELOPMENTAL STAGE: POTENTIALLY SUBJECT TO DEVELOPMENTAL
REGULATION.
CC -1- SIMILARITY: BELONGS TO THE PROTEASOME SUBUNIT S14 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: V00194; CAA23485.1;
CC EMBL: M11429; AAA33223.1;
CC PIR: A03384; QXDO.
CC DictyDb: D001002;
KW Proteasome; Developmental protein.
FT NON_TER 88
SQ SEQUENCE 88 AA; 10130 MW; 370B5997E882B853 CRC64;
Query Match 13.0%; Score 6; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 KTL5LK 8
Db 83 KTL5LK 88

RESULT 11
CYC6_PORPU STANDARD; PRT; 110 AA.
ID CYC6_PORPU
AC P31200;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytochrome C6 precursor (Soluble cytochrome F) (Cytochrome C553).
GN PETJ.

OS Porphyra purpurea.
OG Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2787;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AVONPORT;
RA Reith M.E., Munnholland J.;
RT "Complete nucleotide sequence of the Porphyra purpurea chloroplast
genome."
RL Plant Mol. Biol. Rep. 13:333-335(1995).
CC -1- FUNCTION: CYTOCHROME C6 IS A MONOMERIC MONOMER. IT FUNCTIONS AS AN
ELECTRON CARRIER BETWEEN MEMBRANE-BOUND CYTOCHROME F AND P700 IN
THE PHOTOPHOSPHORYLATION CHAIN IN CHLOROPLASTS AND ALGAE. IT
SUBSTITUTES FOR PLASTOCYANIN IN COPPER-DEFICIENT BLUE-GREEN ALGAE
AND IN THE CHLOROPLASTS OF SOME EUKARYOTE ALGAE.
CC -1- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: U38804; AAC08086.1;
CC HSP: P56534; I66S.
DR InterPro: IPR000345; CytC_heme_bind.
DR InterPro: IPR003088; Cyt-CI.
DR InterPro: IPR002329; Cyt-CIC.
DR Pfam: PF00034; cytochrome_c; 1.
DR PRINTS: PR00605; CYTOCHROME_C.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
KW Electron transport; Chloroplast; Photosynthesis; Heme; Thylakoid.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 110 CYTOCHROME C6.
FT BINDING 39 39 HEME (COVALENT).
FT BINDING 42 42 HEME (COVALENT).
FT METAL 43 43 IRON (HEME AXIAL LIGAND).
FT METAL 83 83 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 110 AA; 11745 MW; E94847B79168995A CRC64;
Query Match 13.0%; Score 6; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKKTL5 6
Db 1 MKKTL5 6

RESULT 12
HV3B_HUMAN STANDARD; PRT; 114 AA.
ID HV3B_HUMAN
AC P01763;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region WEA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=83273707; PubMed=6410398;
RA Goni F., Frangione B.;
RT "Amino acid sequence of the Fv region of a human monoclonal IgM
(protein WEA) with antibody activity against 3,4-pyruvylated
galactose in Klebsiella polysaccharides K30 and K33."
RT Proc. Natl. Acad. Sci. U.S.A. 80:4837-4841(1983).
RL

CC -!- MISCELLANEOUS: THIS CHAIN WAS OBTAINED FROM A MONOCLONAL ANTIBODY
CC AGAINST 3,4-PYRUVYLATED GALACTOSE AND ISOLATED FROM A PATIENT WITH
CC WALDENSTROM'S MACROGLOBULINEMIA.

DR PIR: A02046; M3HUWE.
DR HSSP; P01772; 2IG2.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT MOD_RES 114 114
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12256 MW; D88294FF418A07B7 CRC64;

Query Match 13.0%; Score 6; DB 1; Length 114;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 IGGSGS 28
Db 51 IGGSGS 56

RESULT 13

HV3A_HUMAN
ID HV3A_HUMAN STANDARD; PRT; 122 AA.
AC P01762;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V-III region TRO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE (MVELOMA PROTEIN TRO).
RX MEDLINE-76023781; PubMed-809331;
RA Kratzin H., Altevogt P., Ruban E., Kortt A., Starosick K.,
RA Hilschmann N.;
RT "The primary structure of a monoclonal IgA-immunoglobulin (IgA Tro.);
RT II. The amino acid sequence of the H-chain, alpha-type, subgroup III;
RT structure of the complete IgA-molecule."
RL Hoppe-Seyler's Z. Physiol. Chem. 356:1337-1342(1975).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE C REGION IS ALSO GIVEN.

DR PIR: A02045; ALHUTR.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IgV; 1.
KW Immunoglobulin V region.
FT MOD_RES 122 122
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 13472 MW; 2E21A11DA04D80F9 CRC64;

Query Match 13.0%; Score 6; DB 1; Length 122;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 IGGSGS 28
Db 51 IGGSGS 56

RESULT 14

RADC_COXBU
ID RADC_COXBU STANDARD; PRT; 142 AA.
AC O85403;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA repair protein radC (Fragment).
GN RADC.

OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
OC Coxiella group; Coxiella.
OX NCBI_TaxID=777;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE PHASE I;
RA Williams H., Jaeger C., Baljer G.;
RT "Physical and genetic map of the obligate intracellular bacterium
RT Coxiella burnetii."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE RADC FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC EMBL; AF064960; AAD09944.1;
DR InterPro: IPR001405; RadC.
DR ProDom: PD007415; RadC; 1.
KW DNA repair.
FT NON_TER 1 1

SQ SEQUENCE 142 AA; 15295 MW; 8C621E49EBF54D43 CRC64;
Query Match 13.0%; Score 6; DB 1; Length 142;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 IIGGS 26
Db 119 IIGGS 124

RESULT 15

Y805_AQUAE
ID Y805_AQUAE STANDARD; PRT; 155 AA;
AC O66989;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_805.
GN AQ_805.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE-98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Shear M.A., Keiler M., Auway M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

CC EMBL; AF000708; AAC06954.1;
DR Hypothetical protein: Complete proteome.
KW Hypothetical protein: Complete proteome.
SQ SEQUENCE 155 AA; 18152 MW; 16933A2AC7344531 CRC64;

Query Match 13.0%; Score 6; DB 1; Length 155;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 LFNRSF 40
|||||
Db 54 LFNRSF 59

Search completed: November 5, 2002, 11:05:25
Job time : 57.3582 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:58:27 ; Search time 30.209 seconds
(without alignments)
263.424 Million cell updates/sec

Title: US-09-833-017B-2
Perfect score: 46
Sequence: 1 MKKTLKLNDFKEIKTDELE.....GSLSITFLNRSFTQALCK 46

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL19.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organella.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	46	2	Q99QI5
2	41	89.1	46	2	Q9APK7
3	27	58.7	43	2	Q9APK6
4	9	19.6	363	2	Q9R7H6
5	7	15.2	88	16	Q98RG2
6	7	15.2	96	12	Q83163
7	7	15.2	239	17	Q96Y40
8	7	15.2	240	16	Q9I488
9	7	15.2	240	15	Q92WS3
10	7	15.2	253	13	Q90315
11	7	15.2	253	16	Q67894
12	7	15.2	268	16	Q98CC9
13	7	15.2	276	10	Q9SEW3
14	7	15.2	284	15	Q92S62
15	7	15.2	287	13	Q90314
16	7	15.2	346	13	Q90WN4

17	7	15.2	370	10	Q9LYK0
18	7	15.2	411	4	Q14422
19	7	15.2	437	16	Q92EV2
20	7	15.2	493	10	Q9L785
21	7	15.2	494	10	Q9FXA1
22	7	15.2	579	4	Q9UEI7
23	7	15.2	612	12	Q9J5F6
24	7	15.2	618	10	Q23121
25	7	15.2	628	5	Q960L0
26	7	15.2	767	5	Q9V926
27	7	15.2	808	16	Q9JVK8
28	7	15.2	929	5	Q9VR32
29	7	15.2	1183	16	Q9ROH9
30	7	15.2	1358	5	Q17892
31	7	15.2	1395	4	Q96SB7
32	7	15.2	2771	5	Q9VLS3
33	7	15.2	3972	16	P73139
34	6	13.0	48	16	Q9PBI2
35	6	13.0	51	2	Q9XB45
36	6	13.0	51	12	Q9IG53
37	6	13.0	57	10	Q93998
38	6	13.0	63	6	Q9XSB0
39	6	13.0	69	12	Q41074
40	6	13.0	69	12	Q91PP4
41	6	13.0	97	10	Q9C7H0
42	6	13.0	102	17	Q98597
43	6	13.0	104	16	Q9JY34
44	6	13.0	105	5	Q9U514
45	6	13.0	114	17	Q96Z48
46	6	13.0	115	5	Q17259
47	6	13.0	119	2	Q9RAF5
48	6	13.0	120	2	Q9ALF6
49	6	13.0	120	2	Q9ALF2
50	6	13.0	120	2	Q9ALB0
51	6	13.0	121	2	Q9ALF4
52	6	13.0	121	2	Q9ALC3
53	6	13.0	121	17	Q97AW7
54	6	13.0	122	16	Q9RU75
55	6	13.0	127	16	Q9KUT9
56	6	13.0	133	16	Q9A2R6
57	6	13.0	134	12	Q84484
58	6	13.0	136	2	Q9RC70
59	6	13.0	140	4	Q9BSC0
60	6	13.0	141	13	Q57333
61	6	13.0	141	16	Q9JS86
62	6	13.0	145	17	Q976P9
63	6	13.0	152	5	Q98596
64	6	13.0	152	11	Q9D2A9
65	6	13.0	154	15	Q9CHW9
66	6	13.0	156	11	Q99MF7
67	6	13.0	157	2	Q9JMX4
68	6	13.0	167	2	Q9AKX7
69	6	13.0	167	10	Q9M6N1
70	6	13.0	169	5	Q9NFN1
71	6	13.0	177	15	Q9QRM3
72	6	13.0	180	16	Q98FX0
73	6	13.0	181	15	Q901F3
74	6	13.0	186	10	Q04571
75	6	13.0	186	11	Q62146
76	6	13.0	189	2	Q9KXW7
77	6	13.0	189	11	Q9CAV1
78	6	13.0	191	11	Q9D0T7
79	6	13.0	196	2	Q93KL7
80	6	13.0	200	4	Q9BR12
81	6	13.0	200	5	Q16314
82	6	13.0	200	6	Q95LH6
83	6	13.0	201	15	Q83184
84	6	13.0	201	16	Q92R37
85	6	13.0	202	15	Q9JBC7
86	6	13.0	202	16	Q9PLC4
87	6	13.0	203	16	Q97R00
88	6	13.0	204	11	Q9EPN1
89	6	13.0	205	5	Q15772

Q9LYK0	arabidopsis
Q14422	homo sapien
Q92EV2	listeria in
Q9L785	arabidopsis
Q9FXA1	arabidopsis
Q9UEI7	homo sapien
Q9J5F6	fowlpox vir
Q23121	arabidopsis
Q960L0	drosohilla
Q9V926	drosohilla
Q9JVK8	neisseria m
Q9VR32	drosohilla
Q9ROH9	mycoplasma
Q17892	caenorhabdi
Q96SB7	homo sapien
Q9VLS3	drosohilla
P73139	synecocyst
Q9PBI2	xytella fas
Q9XB45	escherichia
Q9IG53	chilo iride
Q93998	betula verr
Q9XSB0	bos taurus
Q41074	paramecium
Q91PP4	hepatitis b
Q9C7H0	arabidopsis
Q98597	pyrococcus
Q9JY34	neisseria m
Q9U514	manduca sex
Q96Z48	sulfolobus
Q17259	brachionus
Q9RAF5	raistonia s
Q9ALF6	uncultured
Q9ALF2	uncultured
Q9ALB0	uncultured
Q9ALF4	uncultured
Q9ALC3	uncultured
Q97AW7	thermoplasm
Q9RU75	deinococcus
Q9KUT9	vibrio chol
Q9A2R6	caulobacter
Q84484	paramecium
Q9RC70	bacillus ha
Q9BSC0	homo sapien
Q57333	brachydanio
Q9JS86	chlamydia p
Q976P9	sulfolobus
Q98596	tenebrio mo
Q9D2A9	mus musculus
Q9CHW9	lactococcus
Q99MF7	rattus norv
Q9JMX4	listeria mo
Q9AKX7	legionella
Q9M6N1	cicer ariet
Q9NFN1	schistosoma
Q9QRM3	human immun
Q98FX0	rhizobium l
Q901F3	human immun
Q04571	arabidopsis
Q62146	mus musculus
Q9KXW7	streptomyce
Q9CAV1	mus musculus
Q9D0T7	mus musculus
Q93KL7	uncultured
Q9BR12	homo sapien
Q16314	caenorhabdi
Q95LH6	bos taurus
Q83184	mink cell f
Q92R37	rhizobium m
Q9JBC7	human immun
Q9PLC4	chlamydia m
Q97R00	streptococc
Q9EPN1	rattus norv
Q15772	trypanosoma

90	6	13.0	205	5	015777	015777 trypanosoma	163	6	13.0	306	17	Q971C2	Q971c2 sulfolobus
91	6	13.0	206	2	Q9AF58	Q9af58 desulfotoma	164	6	13.0	308	11	Q9CVZ7	Q9cvz7 mus musculus
92	6	13.0	206	2	Q93TR6	Q93tr6 uncultured	165	6	13.0	309	17	Q9UYG8	Q9uyg8 pyrococcus
93	6	13.0	208	16	Q9CPJ3	Q9cpj3 pasteurella	166	6	13.0	310	10	Q9ZTJ0	Q9ztj0 zea mays (m
94	6	13.0	213	5	Q60953	Q60953 dictyosteli	167	6	13.0	312	5	Q62365	Q62365 caenorhabdi
95	6	13.0	213	13	Q12993	Q12993 oryza sat	168	6	13.0	313	16	Q98E28	Q98e28 rhizobium l
96	6	13.0	213	16	Q84471	Q84471 chlamydia t	169	6	13.0	314	1	P94510	P94510 methanosarc
97	6	13.0	215	13	Q12999	Q12999 xenopus lae	170	6	13.0	314	2	Q9L262	Q9l262 streptomyce
98	6	13.0	217	13	Q9PT27	Q9pt27 oncorhynch	171	6	13.0	317	5	Q9VBH7	Q9vbh7 drosophila
99	6	13.0	217	13	Q9PUS6	Q9pus6 brachydanio	172	6	13.0	318	2	Q93F56	Q93f56 pseudomonas
100	6	13.0	217	13	Q9PDES0	Q9des0 fugu rubrip	173	6	13.0	318	4	Q9UMU8	Q9umu8 homo sapien
101	6	13.0	218	13	Q9PUS3	Q9pus3 brachydanio	174	6	13.0	318	16	Q987Q7	Q987q7 rhizobium l
102	6	13.0	220	4	Q96R14	Q96r14 homo sapien	175	6	13.0	319	5	Q43950	Q43950 theileria a
103	6	13.0	220	16	Q25877	Q25877 helicobacte	176	6	13.0	320	2	Q93F58	Q93f58 pseudomonas
104	6	13.0	220	16	Q9ZTJ8	Q9ztj8 helicobacte	177	6	13.0	321	11	Q923C4	Q923c4 mus musculus
105	6	13.0	224	2	Q9XG66	Q9xg66 streptomyce	178	6	13.0	322	17	Q97CK7	Q97ck7 thermoplasm
106	6	13.0	224	5	Q9V7H8	Q9v7h8 drosophila	179	6	13.0	323	2	Q937X6	Q937x6 edwardsiell
107	6	13.0	224	5	Q955A6	Q955a6 drosophila	180	6	13.0	323	10	Q98SB0	Q98sd0 guillardi
108	6	13.0	227	2	Q9XJ34	Q9xj34 pseudomonas	181	6	13.0	326	16	P74469	P74469 synechocyst
109	6	13.0	227	16	Q51441	Q51441 pseudomonas	182	6	13.0	329	16	Q83509	Q83509 treponema p
110	6	13.0	229	5	Q62545	Q62545 hydra atten	183	6	13.0	330	17	Q9UYJ6	Q9uyj6 pyrococcus
111	6	13.0	229	10	Q41723	Q41723 zinnia eleg	184	6	13.0	331	3	Q9HF69	Q9hf69 candida dub
112	6	13.0	230	13	Q13000	Q13000 xenopus lae	185	6	13.0	336	2	Q87519	Q87519 escherichia
113	6	13.0	232	16	Q98QM4	Q98qm4 mycoplasma	186	6	13.0	340	3	P78846	P78846 schizosacch
114	6	13.0	233	10	Q23716	Q23716 arabidopsis	187	6	13.0	342	2	Q924N6	Q924n6 pasteurella
115	6	13.0	234	17	Q9HSL9	Q9hsl9 halobacteri	188	6	13.0	342	5	Q9VKR8	Q9vkr8 drosophila
116	6	13.0	235	16	Q9RR32	Q9rr32 vibrio chol	189	6	13.0	342	5	Q9GRL0	Q9grl0 leishmania
117	6	13.0	236	3	Q9PVJ3	Q9pvj3 schizosacch	190	6	13.0	345	4	Q96IH2	Q96ih2 homo sapien
118	6	13.0	238	2	P94902	P94902 lactobacilli	191	6	13.0	349	5	P91081	P91081 caenorhabdi
119	6	13.0	239	4	Q96J55	Q96j55 homo sapien	192	6	13.0	350	3	Q96W98	Q96w98 candida alb
120	6	13.0	243	16	Q928R7	Q928r7 chlamydia p	193	6	13.0	352	2	Q93F60	Q93f60 pseudomonas
121	6	13.0	243	16	Q930D1	Q930d1 rhizobium m	194	6	13.0	352	2	Q93F57	Q93f57 pseudomonas
122	6	13.0	244	5	Q20804	Q20804 caenorhabdi	195	6	13.0	352	2	Q93F48	Q93f48 pseudomonas
123	6	13.0	246	5	Q18894	Q18894 caenorhabdi	196	6	13.0	353	2	Q93F51	Q93f51 pseudomonas
124	6	13.0	246	10	Q9LST9	Q9lst9 oryza sativ	197	6	13.0	354	2	Q93F52	Q93f52 pseudomonas
125	6	13.0	247	17	Q979U1	Q979u1 thermoplasm	198	6	13.0	354	16	Q25675	Q25675 helicobacte
126	6	13.0	248	17	Q9HKF9	Q9hkf9 thermoplasm	199	6	13.0	354	16	Q92M30	Q92m30 helicobacte
127	6	13.0	249	2	Q9R3A5	Q9r3a5 clostridium	200	6	13.0	359	3	Q9V7Q2	Q9v7q2 drosophila
128	6	13.0	250	2	Q9Z4F5	Q9z4f5 shigella so	201	6	13.0	360	3	Q00296	Q00296 beaueveria b
129	6	13.0	255	4	Q96BD8	Q96bd8 homo sapien	202	6	13.0	360	5	Q9VXL2	Q9vxl2 drosophila
130	6	13.0	256	12	Q36455	Q36455 chino del t	203	6	13.0	360	16	Q928L6	Q928l6 chlamydia p
131	6	13.0	256	16	Q9A270	Q9a270 caulobacter	204	6	13.0	366	16	Q99TT7	Q99tt7 staphylococ
132	6	13.0	257	17	Q97BL7	Q97bl7 thermoplasm	205	6	13.0	368	2	Q93GD2	Q93gd2 desulfotoma
133	6	13.0	261	17	Q9HL98	Q9hl98 thermoplasm	206	6	13.0	369	2	Q93GE2	Q93ge2 desulfotoma
134	6	13.0	262	16	Q9PQEO	Q9pqeo ureaplasma	207	6	13.0	370	5	Q95TX4	Q95tx4 drosophila
135	6	13.0	265	10	Q22782	Q22782 arabidopsis	208	6	13.0	370	16	Q97PU9	Q97pu9 streptococc
136	6	13.0	270	17	Q97W94	Q97w94 sulfolobus	209	6	13.0	372	2	Q57178	Q57178 salmonella
137	6	13.0	271	16	Q9H027	Q9h027 homo sapien	210	6	13.0	372	2	Q54034	Q54034 salmonella
138	6	13.0	271	16	Q9K8F9	Q9k8f9 bacillus ha	211	6	13.0	372	2	Q54036	Q54036 salmonella
139	6	13.0	271	17	Q975C3	Q975c3 sulfolobus	212	6	13.0	372	2	Q54025	Q54025 salmonella
140	6	13.0	274	16	Q9K9J7	Q9k9j7 bacillus ha	213	6	13.0	372	2	Q54026	Q54026 salmonella
141	6	13.0	274	16	Q928Z8	Q928z8 listeria in	214	6	13.0	372	2	Q54027	Q54027 salmonella
142	6	13.0	276	5	Q9W5C7	Q9w5c7 drosophila	215	6	13.0	372	2	Q54028	Q54028 salmonella
143	6	13.0	277	16	Q66839	Q66839 aquifex aeo	216	6	13.0	372	2	Q54029	Q54029 salmonella
144	6	13.0	278	16	Q97JB3	Q97jb3 clostridium	217	6	13.0	372	2	Q54030	Q54030 salmonella
145	6	13.0	285	5	Q94477	Q94477 dictyosteli	218	6	13.0	372	2	Q54031	Q54031 salmonella
146	6	13.0	285	16	Q9KDP6	Q9kdp6 bacillus ha	219	6	13.0	372	2	Q54032	Q54032 salmonella
147	6	13.0	287	16	Q97F64	Q97f64 clostridium	220	6	13.0	372	2	Q54037	Q54037 salmonella
148	6	13.0	288	16	Q9PB74	Q9pb74 xylella fas	221	6	13.0	372	2	Q54038	Q54038 salmonella
149	6	13.0	289	5	Q9V813	Q9v813 drosophila	222	6	13.0	372	2	Q54039	Q54039 salmonella
150	6	13.0	289	5	Q01509	Q01509 caenorhabdi	223	6	13.0	372	2	Q54040	Q54040 salmonella
151	6	13.0	294	2	Q93F53	Q93f53 pseudomonas	224	6	13.0	372	2	Q56052	Q56052 salmonella
152	6	13.0	294	17	Q97Z86	Q97z86 sulfolobus	225	6	13.0	375	2	Q93GE0	Q93ge0 desulfotoma
153	6	13.0	297	11	Q64106	Q64106 rattus sp.	226	6	13.0	375	10	Q9SVA0	Q9sva0 arabidopsis
154	6	13.0	298	16	Q9KC62	Q9kc62 bacillus ha	227	6	13.0	377	3	Q96VM1	Q96vm1 beaueveria b
155	6	13.0	300	2	Q9RJ57	Q9rj57 streptomyce	228	6	13.0	377	4	Q96RQ1	Q96rq1 homo sapien
156	6	13.0	300	2	Q93UN6	Q93un6 vibrio vuln	229	6	13.0	378	4	Q9NZA3	Q9nza3 fusarium sp
157	6	13.0	300	16	Q91259	Q91259 pseudomonas	230	6	13.0	379	3	Q9V759	Q9v759 beaueveria b
158	6	13.0	304	17	Q58240	Q58240 pyrococcus	231	6	13.0	381	2	Q9XC30	Q9xc30 acinetobact
159	6	13.0	305	16	Q9P8R5	Q9p8r5 ureaplasma	232	6	13.0	381	10	Q9CAQ3	Q9caq3 arabidopsis
160	6	13.0	306	2	P71182	P71182 enterobacte	233	6	13.0	386	6	Q9BGU5	Q9bgu5 bos taurus
161	6	13.0	306	2	Q9AHH2	Q9anh2 comamonas t	234	6	13.0	387	10	Q9FN48	Q9fn48 arabidopsis
162	6	13.0	306	17	Q9V055	Q9v055 pyrococcus	235	6	13.0				

236	6	13.0	392	5	Q9V313	Q9V313 drosophila	309	6	13.0	559	4	Q9NVH1	Q9NVH1 homo sapien
237	6	13.0	394	10	Q9AJUB	Q9AJUB arabidopsis	310	6	13.0	560	16	Q25362	Q25362 helicobacte
238	6	13.0	396	13	Q93428	Q93428 chionodraco	311	6	13.0	560	16	Q92LI9	Q92LI9 helicobacte
239	6	13.0	396	13	Q9DEX3	Q9DEX3 clupea hare	312	6	13.0	561	2	Q9EYF3	Q9EYF3 escherichia
240	6	13.0	398	3	Q00070	Q00070 aspergillus	313	6	13.0	561	5	Q76134	Q76134 tetrahymena
241	6	13.0	398	13	P87370	P87370 conorhynchu	314	6	13.0	563	4	Q96F26	Q96F26 homo sapien
242	6	13.0	399	3	Q9P8V4	Q9P8V4 coccidioides	315	6	13.0	564	2	Q9L900	Q9L900 enterococcu
243	6	13.0	399	13	Q9DD89	Q9DD89 bractydanio	316	6	13.0	567	10	Q64843	Q64843 arabidopsis
244	6	13.0	404	5	Q9V717	Q9V717 drosophila	317	6	13.0	569	16	Q9PHX1	Q9PHX1 campylobact
245	6	13.0	415	12	Q9LFF5	Q9LFF5 chilo iride	318	6	13.0	571	16	Q9I502	Q9I502 pseudomonas
246	6	13.0	416	16	Q9ZEB0	Q9ZEB0 listeria in	319	6	13.0	575	5	Q45066	Q45066 caenorhabdi
247	6	13.0	417	1	Q93650	Q93650 archaeoglob	320	6	13.0	579	10	Q9M056	Q9M056 arabidopsis
248	6	13.0	418	16	Q9PP65	Q9PP65 campylobact	321	6	13.0	580	5	Q9V480	Q9V480 drosophila
249	6	13.0	419	12	Q9QD05	Q9QD05 tobacco vel	322	6	13.0	583	5	Q61711	Q61711 caenorhabdi
250	6	13.0	421	16	Q92WM5	Q92WM5 rhizobium m	323	6	13.0	583	5	Q27472	Q27472 caenorhabdi
251	6	13.0	422	5	Q76607	Q76607 caenorhabdi	324	6	13.0	583	10	Q9L449	Q9L449 arabidopsis
252	6	13.0	423	5	Q9VWP7	Q9VWP7 drosophila	325	6	13.0	590	11	Q88337	Q88337 rattus norv
253	6	13.0	425	10	Q9WA53	Q9WA53 arabidopsis	326	6	13.0	593	16	Q92HB9	Q92HB9 rickettsia
254	6	13.0	427	16	Q9R0N5	Q9R0N5 deinococcus	327	6	13.0	594	16	Q92CY6	Q92CY6 rickettsia
255	6	13.0	429	3	Q6G103	Q6G103 saccharomyc	328	6	13.0	597	2	P72416	P72416 streptococ
256	6	13.0	433	16	Q97GDS	Q97GDS clostridium	329	6	13.0	598	2	Q9JMY5	Q9JMY5 listeria mo
257	6	13.0	434	16	Q97K95	Q97K95 clostridium	330	6	13.0	599	5	Q95N20	Q95N20 spodoptera
258	6	13.0	438	11	Q920C3	Q920C3 mus musculus	331	6	13.0	608	4	Q9NQX0	Q9NQX0 homo sapien
259	6	13.0	447	10	Q9SIA8	Q9SIA8 arabidopsis	332	6	13.0	611	5	Q9VYR3	Q9VYR3 drosophila
260	6	13.0	448	2	Q93TD3	Q93TD3 pseudomonas	333	6	13.0	612	3	Q9US12	Q9US12 schizosacch
261	6	13.0	452	3	Q9UVS9	Q9UVS9 aspergillus	334	6	13.0	612	5	Q22779	Q22779 caenorhabdi
262	6	13.0	452	10	Q9LRP8	Q9LRP8 arabidopsis	335	6	13.0	615	10	Q38957	Q38957 arabidopsis
263	6	13.0	454	8	Q35213	Q35213 oenothera b	336	6	13.0	621	10	Q94EX4	Q94EX4 cucumis mel
264	6	13.0	455	2	Q9F160	Q9F160 arabidopsis	337	6	13.0	626	11	Q88810	Q88810 rattus norv
265	6	13.0	457	10	Q9M158	Q9M158 arabidopsis	338	6	13.0	633	2	Q9JMY6	Q9JMY6 listeria mo
266	6	13.0	460	2	Q9X3V2	Q9X3V2 pseudomonas	339	6	13.0	634	16	Q97MH1	Q97MH1 clostridium
267	6	13.0	465	10	Q9PHT1	Q9PHT1 arabidopsis	340	6	13.0	635	4	Q969V5	Q969V5 homo sapien
268	6	13.0	466	4	Q15846	Q15846 homo sapien	341	6	13.0	637	2	Q33838	Q33838 thermotoga
269	6	13.0	466	10	Q93VZ2	Q93VZ2 arabidopsis	342	6	13.0	639	11	Q9CZ33	Q9CZ33 mus musculus
270	6	13.0	466	16	Q99297	Q99297 streptococ	343	6	13.0	639	16	Q9S5X2	Q9S5X2 thermotoga
271	6	13.0	468	10	Q82730	Q82730 arabidopsis	344	6	13.0	643	5	Q3XZ22	Q3XZ22 drosophila
272	6	13.0	471	2	Q9E2G5	Q9E2G5 streptococ	345	6	13.0	649	17	Q980R1	Q980R1 sulfolobus
273	6	13.0	472	3	Q9U082	Q9U082 schizosacch	346	6	13.0	662	8	Q36521	Q36521 plactymonas
274	6	13.0	472	17	Q972D2	Q972D2 sulfolobus	347	6	13.0	671	5	Q9TZK7	Q9TZK7 caenorhabdi
275	6	13.0	476	2	P72120	P72120 pseudomonas	348	6	13.0	673	11	Q9R0Q1	Q9R0Q1 mus musculus
276	6	13.0	478	5	Q9W0H5	Q9W0H5 drosophila	349	6	13.0	678	16	Q983P2	Q983P2 rhizobium l
277	6	13.0	478	17	Q57804	Q57804 pyrococcus	350	6	13.0	688	16	Q92BQ2	Q92BQ2 listeria in
278	6	13.0	479	16	Q97SA7	Q97SA7 streptococ	351	6	13.0	693	10	Q948P4	Q948P4 cucumis mel
279	6	13.0	481	17	Q9V2K8	Q9V2K8 rhodocyclu	352	6	13.0	693	11	Q9VW66	Q9VW66 mus musculus
280	6	13.0	484	2	Q9JP99	Q9JP99 rhodocyclu	353	6	13.0	693	17	Q96XH6	Q96XH6 sulfolobus
281	6	13.0	484	5	Q26475	Q26475 schistocerc	354	6	13.0	697	11	Q9NFO0	Q9NFO0 mus musculus
282	6	13.0	493	16	Q06294	Q06294 mycobacteri	355	6	13.0	698	10	Q9FN17	Q9FN17 arabidopsis
283	6	13.0	495	2	Q9AD76	Q9AD76 streptomyc	356	6	13.0	702	17	Q28289	Q28289 archaeoglob
284	6	13.0	497	4	Q9BXN7	Q9BXN7 homo sapien	357	6	13.0	715	5	Q02140	Q02140 caenorhabdi
285	6	13.0	499	3	Q9C4B1	Q9C4B1 aspergillus	358	6	13.0	721	16	P95097	P95097 mycobacteri
286	6	13.0	499	3	Q12951	Q12951 aspergillus	359	6	13.0	726	16	Q9AAB8	Q9AAB8 caulobacter
287	6	13.0	501	5	Q9V764	Q9V764 drosophila	360	6	13.0	737	2	Q9EYQ5	Q9EYQ5 clostridium
288	6	13.0	504	4	Q9H4D7	Q9H4D7 homo sapien	361	6	13.0	737	5	Q94706	Q94706 physarum po
289	6	13.0	506	3	Q96X02	Q96X02 penicillium	362	6	13.0	781	4	Q95560	Q95560 homo sapien
290	6	13.0	506	3	Q96VA1	Q96VA1 aspergillus	363	6	13.0	781	5	Q960F0	Q960F0 drosophila
291	6	13.0	506	3	Q96VA0	Q96VA0 aspergillus	364	6	13.0	787	5	Q9N5V0	Q9N5V0 caenorhabdi
292	6	13.0	507	4	Q96CL7	Q96CL7 homo sapien	365	6	13.0	791	4	Q9C0B6	Q9C0B6 homo sapien
293	6	13.0	508	5	Q20604	Q20604 caenorhabdi	366	6	13.0	796	5	Q9TVF4	Q9TVF4 caenorhabdi
294	6	13.0	509	11	Q920C2	Q920C2 mus musculus	367	6	13.0	802	13	Q42127	Q42127 xenopus lae
295	6	13.0	510	3	Q74288	Q74288 emericella	368	6	13.0	804	3	Q13768	Q13768 schizosacch
296	6	13.0	511	10	Q942V7	Q942V7 oryza sativ	369	6	13.0	804	10	Q64770	Q64770 arabidopsis
297	6	13.0	519	5	Q22017	Q22017 caenorhabdi	370	6	13.0	805	5	Q62195	Q62195 caenorhabdi
298	6	13.0	519	15	Q74230	Q74230 human immun	371	6	13.0	809	5	Q9Y119	Q9Y119 drosophila
299	6	13.0	521	15	Q76633	Q76633 human immun	372	6	13.0	817	5	Q9XTF3	Q9XTF3 caenorhabdi
300	6	13.0	526	5	Q9V9Y4	Q9V9Y4 drosophila	373	6	13.0	818	5	Q9U434	Q9U434 drosophila
301	6	13.0	528	11	Q9LWI6	Q9LWI6 oryza sativ	374	6	13.0	825	5	Q23484	Q23484 caenorhabdi
302	6	13.0	529	11	Q91V87	Q91V87 mus musculus	375	6	13.0	828	5	Q9VKK7	Q9VKK7 drosophila
303	6	13.0	530	2	Q87219	Q87219 lactococcus	376	6	13.0	829	2	Q44204	Q44204 agrobacteri
304	6	13.0	534	5	Q961B3	Q961B3 drosophila	377	6	13.0	831	10	Q64781	Q64781 arabidopsis
305	6	13.0	540	2	Q68145	Q68145 burkholderi	378	6	13.0	833	2	Q9R6C4	Q9R6C4 agrobacteri
306	6	13.0	548	2	Q93A25	Q93A25 buchnera ap	379	6	13.0	833	2	Q52297	Q52297 agrobacteri
307	6	13.0	553	5	Q9VRZ2	Q9VRZ2 drosophila	380	6	13.0	842	5	Q9U3A8	Q9U3A8 caenorhabdi
308	6	13.0	556	5	Q95025	Q95025 dictyosteli	381	6	13.0	856	10	Q942J7	Q942J7 oryza sativ

382	6	13.0	860	5	Q9NDT9	Q9ndt9 balanus amp	455	6	13.0	1643	11	Q91WX9	Q91wx9 rattus norv
383	6	13.0	862	5	Q9W4P0	Q9w4p0 drosophila	456	6	13.0	1670	5	Q9U572	Q9u572 penaeus mon
384	6	13.0	872	4	Q9H6U7	Q9h6u7 homo sapien	457	6	13.0	1688	13	Q57483	Q57483 rana catesb
385	6	13.0	872	4	Q9G6N3	Q9g6n3 homo sapien	458	6	13.0	1769	4	Q9P273	Q9p273 homo sapien
386	6	13.0	874	10	Q81723	Q81723 arabidopsis	459	6	13.0	1816	12	Q9JAD4	Q9jad4 soil-borne
387	6	13.0	875	17	Q97WS1	Q97ws1 sulfolobus	460	6	13.0	1816	12	Q9JAD3	Q9jad3 soil-borne
388	6	13.0	879	5	Q76978	Q76978 sycon rapha	461	6	13.0	1816	12	Q9JAD2	Q9jad2 soil-borne
389	6	13.0	890	12	Q88171	Q88171 strawberry	462	6	13.0	1816	12	Q9JWB1	Q9jwb1 soil-borne
390	6	13.0	893	16	Q9X1V4	Q9x1v4 thermotoga	463	6	13.0	1816	12	Q9DJG6	Q9djg6 soil-borne
391	6	13.0	909	6	Q9T7T4	Q9t7t4 oryctolagus	464	6	13.0	1816	12	Q9IDN3	Q9idn3 soil-borne
392	6	13.0	913	2	Q9ZEX7	Q9zex7 pseudomonas	465	6	13.0	1828	12	Q89249	Q89249 soil-borne
393	6	13.0	914	2	Q9P6V9	Q9p6v9 alteromonas	466	6	13.0	1855	5	Q9TX75	Q9tx75 plasmodium
394	6	13.0	914	11	Q91W15	Q91w15 mus musculus	467	6	13.0	1855	5	Q9BHN0	Q9bhn0 plasmodium
395	6	13.0	945	17	Q974N4	Q974n4 sulfolobus	468	6	13.0	1858	12	Q9WJD7	Q9wjdt chinese whe
396	6	13.0	955	4	Q9I2W1	Q9i2w1 homo sapien	469	6	13.0	1858	12	Q91QY9	Q91qy9 chinese whe
397	6	13.0	961	10	Q9SXX6	Q9sxn6 nicotiana t	470	6	13.0	1864	10	Q9FYL7	Q9fyl7 arabidopsis
398	6	13.0	968	10	Q04623	Q04623 arabidopsis	471	6	13.0	1876	5	Q95Q08	Q95qq8 caenorhabdi
399	6	13.0	971	5	Q9TYL9	Q9tyl9 caenorhabdi	472	6	13.0	1877	5	Q9XXW1	Q9xxw1 plasmodium
400	6	13.0	977	10	Q9LLI0	Q9lli0 phaseolus l	473	6	13.0	1897	3	Q13428	Q13428 candida alb
401	6	13.0	977	16	Q25308	Q25308 helicobacte	474	6	13.0	1912	4	Q95226	Q95226 homo sapien
402	6	13.0	978	4	Q96PY8	Q96py8 homo sapien	475	6	13.0	1912	5	Q9U0H1	Q9u0h1 plasmodium
403	6	13.0	1010	4	Q15058	Q15058 homo sapien	476	6	13.0	1977	4	Q9UHB1	Q9uhb1 homo sapien
404	6	13.0	1012	10	Q82677	Q82677 chenopodium	477	6	13.0	1981	11	Q92327	Q92327 rattus norv
405	6	13.0	1013	4	Q96PY9	Q96py9 homo sapien	478	6	13.0	1985	11	Q9JIS7	Q9jis7 mus musculus
406	6	13.0	1013	10	Q9LKK3	Q9lkk3 arabidopsis	479	6	13.0	2006	5	Q9VME2	Q9vmf2 drosophila
407	6	13.0	1013	10	Q93ZS5	Q93zs5 arabidopsis	480	6	13.0	2090	3	Q9P8X4	Q9p8x4 neurospora
408	6	13.0	1021	10	Q9LHH8	Q9lhh8 arabidopsis	481	6	13.0	2100	3	P87112	P87112 schizosacch
409	6	13.0	1024	4	Q96RT8	Q96rt8 homo sapien	482	6	13.0	2110	5	Q9VRA6	Q9vra6 drosophila
410	6	13.0	1035	5	Q21079	Q21079 caenorhabdi	483	6	13.0	2168	13	Q9OZA7	Q9oza7 brachydanio
411	6	13.0	1054	10	Q9C698	Q9c698 arabidopsis	484	6	13.0	2169	11	Q92305	Q92305 cavia porce
412	6	13.0	1072	11	Q9DBT7	Q9dbt7 mus musculus	485	6	13.0	2271	16	Q99QY4	Q99qy4 staphylococ
413	6	13.0	1072	16	Q92X83	Q92x83 rhizobium m	486	6	13.0	2322	4	Q92675	Q92675 homo sapien
414	6	13.0	1086	5	Q9VM31	Q9vm31 drosophila	487	6	13.0	2346	11	Q9JLC1	Q9jlc1 mus musculus
415	6	13.0	1088	10	Q9C7C9	Q9c7c9 arabidopsis	488	6	13.0	2425	17	Q28859	Q28859 archaeoglob
416	6	13.0	1110	13	Q91Z55	Q91z55 petromyzon	489	6	13.0	2590	13	Q9W7R4	Q9w7r4 brachydanio
417	6	13.0	1120	5	Q17505	Q17505 bombyx mori	490	6	13.0	2706	5	Q97292	Q97292 plasmodium
418	6	13.0	1126	17	Q976A4	Q976a4 sulfolobus	491	6	13.0	2715	11	Q9WTS6	Q9wts6 mus musculus
419	6	13.0	1145	12	Q9IV56	Q9iv56 olive laten	492	6	13.0	2778	5	Q9V9T6	Q9v9t6 drosophila
420	6	13.0	1160	5	P90935	P90935 caenorhabdi	493	6	13.0	3194	16	Q9ZIM3	Q9zim3 helicobacte
421	6	13.0	1164	16	Q9ZJN3	Q9zjn3 helicobacte	494	6	13.0	3512	5	Q62524	Q62524 chironomus
422	6	13.0	1174	5	Q9VB25	Q9vb25 drosophila	495	6	13.0	3512	5	Q62524	Q62524 chironomus
423	6	13.0	1177	12	Q92611	Q92611 pseudorabie	496	6	13.0	3659	16	Q98LN6	Q98ln6 rhizobium l
424	6	13.0	1184	10	Q22725	Q22725 arabidopsis	497	6	13.0	3734	3	Q9CIG0	Q9cig0 kalliichroma
425	6	13.0	1218	5	Q9V507	Q9v507 drosophila	498	6	13.0	3930	16	Q98E20	Q98e20 rhizobium l
426	6	13.0	1222	5	Q9W9J7	Q9w9j7 drosophila	499	6	13.0	4910	3	Q12019	Q12019 saccharomyc
427	6	13.0	1263	11	Q9GYZ2	Q9gyz2 mus musculus	500	6	13.0	8243	3	Q96554	Q96554 cryptospori
428	6	13.0	1286	5	P90936	P90936 caenorhabdi	501	6	13.0	1288	6	Q18758	Q18758 sus scrofa
429	6	13.0	1289	2	Q9F0G7	Q9f0g7 rickettsia	502	6	13.0	26926	4	Q10466	Q10466 homo sapien
430	6	13.0	1291	4	Q9UFV0	Q9ufv0 homo sapien	503	5	10.9	14	6	Q09061	Q09061 bos taurus
431	6	13.0	1308	12	Q9Q9Q8	Q9q9q8 soil-borne	504	5	10.9	14	11	Q35917	Q35917 rattus norv
432	6	13.0	1308	12	Q9Q9R0	Q9q9r0 soil-borne	505	5	10.9	16	2	Q9R4L0	Q9r4l0 spiroplasma
433	6	13.0	1308	12	Q9Q9O6	Q9q9o6 soil-borne	506	5	10.9	25	6	Q9TRV9	Q9trv9 oryctolagus
434	6	13.0	1308	12	Q9QCE8	Q9qce8 soil-borne	507	5	10.9	26	9	Q38370	Q38370 mycobacteri
435	6	13.0	1308	12	Q9D7G5	Q9d7g5 soil-borne	508	5	10.9	29	11	Q99JY5	Q99jy5 mus musculus
436	6	13.0	1308	12	Q91DN2	Q91dn2 soil-borne	509	5	10.9	35	16	Q50887	Q50887 borrelia bu
437	6	13.0	1320	12	Q66359	Q66359 soil-borne	510	5	10.9	40	6	Q9BDE8	Q9bde8 sus scrofa
438	6	13.0	1335	5	Q17250	Q17250 bombyx mori	511	5	10.9	40	12	Q91FE0	Q91fe0 chilo iride
439	6	13.0	1335	5	Q95PE2	Q95pe2 bombyx mori	512	5	10.9	41	2	Q93IX0	Q93ix0 helicobacte
440	6	13.0	1342	5	Q9VHH2	Q9vhh2 drosophila	513	5	10.9	42	11	P70476	P70476 rattus norv
441	6	13.0	1350	12	Q9WIE3	Q9wie3 chinese whe	514	5	10.9	42	16	Q98040	Q98q40 mycoplasma
442	6	13.0	1350	12	Q91QZ0	Q91qz0 chinese whe	515	5	10.9	46	2	Q51952	Q51952 staphylococ
443	6	13.0	1355	2	Q05700	Q05700 synecococc	516	5	10.9	46	4	Q15456	Q15456 homo sapien
444	6	13.0	1359	4	Q9UPR1	Q9upr1 homo sapien	517	5	10.9	46	12	Q71201	Q71201 bovine aden
445	6	13.0	1372	2	Q54151	Q54151 shigella fl	518	5	10.9	47	16	Q9JS22	Q9js22 neisseria m
446	6	13.0	1373	2	Q9RAA2	Q9raa2 rickettsia	519	5	10.9	48	16	Q25662	Q25662 helicobacte
447	6	13.0	1373	2	Q9AL58	Q9al58 shigella fl	520	5	10.9	49	2	Q9F214	Q9f214 roseateles
448	6	13.0	1429	10	Q94H01	Q94hul oryza sativ	521	5	10.9	49	2	Q9FB90	Q9fb90 helicobacte
449	6	13.0	1475	5	Q9N4G4	Q9n4g4 caenorhabdi	522	5	10.9	49	2	Q9FB89	Q9fb89 helicobacte
450	6	13.0	1523	5	Q9N8U8	Q9n8u8 trypanosoma	523	5	10.9	49	2	Q9FB84	Q9fb84 helicobacte
451	6	13.0	1539	5	Q81068	Q81068 arabidopsis	524	5	10.9	49	2	Q9FB83	Q9fb83 helicobacte
452	6	13.0	1559	13	Q73697	Q73697 fugu rubrip	525	5	10.9	50	2	Q9FB83	Q9fb83 helicobacte
453	6	13.0	1583	10	Q9S7A7	Q9s7a7 oryza sativ	526	5	10.9	50	13	Q9BD22	Q9bd22 micropterus
454	6	13.0	1602	11	Q91W25	Q91w25 mesocricetu	527	5	10.9	50	16	O50723	O50723 borrelia bu

528	5	10.9	51	16	051018	051018 borrelia bu	601	5	10.9	75	15	09QPK4	09QPK4 human immun
529	5	10.9	51	16	09RNP7	09RNP7 vibrio chol	602	5	10.9	75	15	09QPK3	09QPK3 human immun
530	5	10.9	52	6	077483	077483 canis famli	603	5	10.9	75	15	09QPK2	09QPK2 human immun
531	5	10.9	52	16	09RPF0	09RPF0 xylella fas	604	5	10.9	75	15	09QPK1	09QPK1 human immun
532	5	10.9	53	2	0938L5	0938L5 streptococ	605	5	10.9	75	15	09QPK0	09QPK0 human immun
533	5	10.9	53	10	065063	065063 picea maria	606	5	10.9	75	15	09QPK9	09QPK9 human immun
534	5	10.9	53	16	09PDD1	09PDD1 xylella fas	607	5	10.9	75	15	09QPK7	09QPK7 human immun
535	5	10.9	53	16	069194	069194 streptococ	608	5	10.9	75	15	09QPK6	09QPK6 human immun
536	5	10.9	54	5	023882	023882 dictyosteli	609	5	10.9	75	15	09QPK5	09QPK5 human immun
537	5	10.9	54	6	028539	028539 oviss aries	610	5	10.9	75	15	09QPK4	09QPK4 human immun
538	5	10.9	55	8	09XDS5	09XDS5 salmonella	611	5	10.9	75	15	09QPK3	09QPK3 human immun
539	5	10.9	55	8	09TBJ4	09TBJ4 cacomantis	612	5	10.9	75	15	09QPK2	09QPK2 human immun
540	5	10.9	56	13	057575	057575 cynops pyrr	613	5	10.9	75	15	09QPK1	09QPK1 human immun
541	5	10.9	56	16	092RVA	092RVA rhizobium m	614	5	10.9	75	15	09QPK0	09QPK0 human immun
542	5	10.9	57	2	052965	052965 enterococu	615	5	10.9	75	15	09QPK9	09QPK9 human immun
543	5	10.9	57	4	096E83	096E83 homo sapien	616	5	10.9	75	15	09QPK8	09QPK8 human immun
544	5	10.9	58	5	095SA9	095SA9 drosophila	617	5	10.9	75	15	09QPK7	09QPK7 human immun
545	5	10.9	58	16	097JS7	097JS7 clostridium	618	5	10.9	75	15	09QPK6	09QPK6 human immun
546	5	10.9	60	10	041933	041933 arabidopsis	619	5	10.9	75	15	09QPK5	09QPK5 human immun
547	5	10.9	61	5	023877	023877 dictyosteli	620	5	10.9	75	15	09QPK4	09QPK4 human immun
548	5	10.9	61	5	026829	026829 trypanosoma	621	5	10.9	75	15	09QPK3	09QPK3 human immun
549	5	10.9	62	16	P76618	P76618 escherichia	622	5	10.9	75	15	09QPK2	09QPK2 human immun
550	5	10.9	63	9	064074	064074 bacterioph	623	5	10.9	75	15	09QPK1	09QPK1 human immun
551	5	10.9	63	16	031947	031947 bacillus su	624	5	10.9	75	15	09QPK0	09QPK0 human immun
552	5	10.9	64	2	09F1E0	09F1E0 acetobacter	625	5	10.9	75	15	09QPK9	09QPK9 human immun
553	5	10.9	64	2	09ANN8	09ANN8 bradyrhizob	626	5	10.9	75	15	09QPK8	09QPK8 human immun
554	5	10.9	64	10	09LNS5	09LNS5 arabidopsis	627	5	10.9	75	15	09QPK7	09QPK7 human immun
555	5	10.9	65	5	023878	023878 dictyosteli	628	5	10.9	75	15	09QPK6	09QPK6 human immun
556	5	10.9	66	5	023609	023609 caenorhabdi	629	5	10.9	75	15	09QPK5	09QPK5 human immun
557	5	10.9	67	8	020149	020149 chlorella v	630	5	10.9	75	15	09QPK4	09QPK4 human immun
558	5	10.9	67	11	09ER13	09ER13 rattus norv	631	5	10.9	75	15	09QPK3	09QPK3 human immun
559	5	10.9	67	16	09JZM3	09JZM3 neisseria m	632	5	10.9	75	15	09QPK2	09QPK2 human immun
560	5	10.9	67	16	09JY08	09JY08 neisseria m	633	5	10.9	75	15	09QPK1	09QPK1 human immun
561	5	10.9	67	16	09CF38	09CF38 lactococcus	634	5	10.9	75	15	09QPK0	09QPK0 human immun
562	5	10.9	68	7	019350	019350 equus cabal	635	5	10.9	75	15	09QPK9	09QPK9 human immun
563	5	10.9	68	8	048131	048131 phrynosoma	636	5	10.9	75	15	09QPK8	09QPK8 human immun
564	5	10.9	69	12	083161	083161 cattulflower	637	5	10.9	75	15	09QPK7	09QPK7 human immun
565	5	10.9	70	4	09H4S1	09H4S1 homo sapien	638	5	10.9	75	15	09QPK6	09QPK6 human immun
566	5	10.9	70	16	09HH86	09HH86 rhizobium l	639	5	10.9	75	15	09QPK5	09QPK5 human immun
567	5	10.9	71	9	038564	038564 bacterioph	640	5	10.9	75	15	09QPK4	09QPK4 human immun
568	5	10.9	71	11	09J177	09J177 mus musculus	641	5	10.9	75	15	09QPK3	09QPK3 human immun
569	5	10.9	71	17	0978M8	0978M8 thermoplasm	642	5	10.9	75	15	09QPK2	09QPK2 human immun
570	5	10.9	72	2	044537	044537 azotobacter	643	5	10.9	75	15	09QPK1	09QPK1 human immun
571	5	10.9	72	5	09TW80	09TW80 drosophila	644	5	10.9	75	15	09QPK0	09QPK0 human immun
572	5	10.9	72	5	09NT13	09NT13 plasmodium	645	5	10.9	75	15	09QPK9	09QPK9 human immun
573	5	10.9	72	11	061682	061682 mus musculus	646	5	10.9	75	15	09QPK8	09QPK8 human immun
574	5	10.9	72	16	074009	074009 synchocyst	647	5	10.9	75	15	09QPK7	09QPK7 human immun
575	5	10.9	74	2	032312	032312 bacillus th	648	5	10.9	75	15	09QPK6	09QPK6 human immun
576	5	10.9	74	5	095013	095013 plasmodium	649	5	10.9	75	15	09QPK5	09QPK5 human immun
577	5	10.9	74	12	09DW87	09DW87 rat cytomeg	650	5	10.9	75	15	09QPK4	09QPK4 human immun
578	5	10.9	75	11	09WUN3	09WUN3 rattus norv	651	5	10.9	75	15	09QPK3	09QPK3 human immun
579	5	10.9	75	15	09QPM9	09QPM9 human immun	652	5	10.9	75	15	09QPK2	09QPK2 human immun
580	5	10.9	75	15	09QPM8	09QPM8 human immun	653	5	10.9	75	15	09QPK1	09QPK1 human immun
581	5	10.9	75	15	09QPM7	09QPM7 human immun	654	5	10.9	75	15	09QPK0	09QPK0 human immun
582	5	10.9	75	15	09QPM5	09QPM5 human immun	655	5	10.9	75	15	09QPK9	09QPK9 human immun
583	5	10.9	75	15	09QPM3	09QPM3 human immun	656	5	10.9	75	15	09QPK8	09QPK8 human immun
584	5	10.9	75	15	09QPM2	09QPM2 human immun	657	5	10.9	75	15	09QPK7	09QPK7 human immun
585	5	10.9	75	15	09QPM1	09QPM1 human immun	658	5	10.9	75	15	09QPK6	09QPK6 human immun
586	5	10.9	75	15	09QPM0	09QPM0 human immun	659	5	10.9	75	15	09QPK5	09QPK5 human immun
587	5	10.9	75	15	09QPL9	09QPL9 human immun	660	5	10.9	75	15	09QPK4	09QPK4 human immun
588	5	10.9	75	15	09QPL8	09QPL8 human immun	661	5	10.9	75	15	09QPK3	09QPK3 human immun
589	5	10.9	75	15	09QPL7	09QPL7 human immun	662	5	10.9	75	15	09QPK2	09QPK2 human immun
590	5	10.9	75	15	09QPL6	09QPL6 human immun	663	5	10.9	75	15	09QPK1	09QPK1 human immun
591	5	10.9	75	15	09QPL5	09QPL5 human immun	664	5	10.9	75	15	09QPK0	09QPK0 human immun
592	5	10.9	75	15	09QPL4	09QPL4 human immun	665	5	10.9	75	15	09QPK9	09QPK9 human immun
593	5	10.9	75	15	09QPL3	09QPL3 human immun	666	5	10.9	75	15	09QPK8	09QPK8 human immun
594	5	10.9	75	15	09QPL2	09QPL2 human immun	667	5	10.9	75	15	09QPK7	09QPK7 human immun
595	5	10.9	75	15	09QPL1	09QPL1 human immun	668	5	10.9	75	15	09QPK6	09QPK6 human immun
596	5	10.9	75	15	09QPL0	09QPL0 human immun	669	5	10.9	75	15	09QPK5	09QPK5 human immun
597	5	10.9	75	15	09QPK9	09QPK9 human immun	670	5	10.9	75	15	09QPK4	09QPK4 human immun
598	5	10.9	75	15	09QPK8	09QPK8 human immun	671	5	10.9	75	15	09QPK3	09QPK3 human immun
599	5	10.9	75	15	09QPK7	09QPK7 human immun	672	5	10.9	75	15	09QPK2	09QPK2 human immun
600	5	10.9	75	15	09QPK6	09QPK6 human immun	673	5	10.9	75	15	09QPK1	09QPK1 human immun

674	77	2	Q9ZGZ1	10.9	747	5	10.9	97	16	Q97S59	Q97S59 streptococc
675	77	2	Q93MD7	10.9	748	5	10.9	98	9	O48467	O48467 bacterioph
676	77	11	O08631	10.9	749	5	10.9	98	10	Q9ZTI9	Q9ZTI9 zea mays (m
677	78	2	Q9ZHK9	10.9	750	5	10.9	98	12	Q9J8B4	Q9J8B4 spodoptera
678	78	12	Q9ZHP9	10.9	751	5	10.9	98	15	Q9YWS7	Q9YWS7 human immun
679	78	16	Q9ZM05	10.9	752	5	10.9	98	16	Q99YF9	Q99YF9 streptococc
680	79	11	Q9Z206	10.9	753	5	10.9	99	2	O4AFU5	O4AFU5 shigella fl
681	79	16	Q9PDY5	10.9	754	5	10.9	99	5	Q9U771	Q9U771 plasmodium
682	80	5	Q9VN17	10.9	755	5	10.9	99	15	O10115	O10115 human immun
683	80	10	Q93483	10.9	756	5	10.9	99	16	Q9CBF0	Q9CBF0 mycobacteri
684	81	11	Q9RON2	10.9	757	5	10.9	99	16	Q92T24	Q92T24 rhizobium m
685	81	16	Q97HRO	10.9	758	5	10.9	100	5	Q9XYF3	Q9XYF3 drosophila
686	82	2	Q93MT9	10.9	759	5	10.9	100	5	Q9XYF2	Q9XYF2 drosophila
687	82	12	Q69144	10.9	760	5	10.9	100	5	Q9XYF1	Q9XYF1 drosophila
688	82	12	O41082	10.9	761	5	10.9	100	5	Q9XYF0	Q9XYF0 drosophila
689	82	16	Q98G26	10.9	762	5	10.9	100	5	Q9XYE9	Q9XYE9 drosophila
690	83	5	Q9BMC1	10.9	763	5	10.9	100	5	Q9XYE8	Q9XYE8 drosophila
691	83	5	Q93875	10.9	764	5	10.9	100	5	Q9XYE7	Q9XYE7 drosophila
692	83	16	Q97M00	10.9	765	5	10.9	100	5	Q9XYE6	Q9XYE6 drosophila
693	83	16	Q92Q95	10.9	766	5	10.9	100	5	Q9XYE5	Q9XYE5 drosophila
694	84	16	Q9K8T4	10.9	767	5	10.9	100	5	Q9XYE4	Q9XYE4 drosophila
695	84	16	Q984M9	10.9	768	5	10.9	100	5	Q9XYE3	Q9XYE3 drosophila
696	84	17	Q97UP8	10.9	769	5	10.9	100	5	Q9XYE2	Q9XYE2 drosophila
697	85	2	O49461	10.9	770	5	10.9	100	5	Q9XYE1	Q9XYE1 drosophila
698	85	4	Q9DE33	10.9	771	5	10.9	100	5	Q9XYE0	Q9XYE0 drosophila
699	85	5	O43364	10.9	772	5	10.9	100	5	Q9TYD9	Q9TYD9 styela clav
700	85	16	Q98MY3	10.9	773	5	10.9	100	10	Q93X17	Q93X17 pharbitis n
701	86	8	Q98XK2	10.9	774	5	10.9	100	16	Q9A7J8	Q9A7J8 caulobacter
702	86	8	Q95807	10.9	775	5	10.9	101	11	O88540	O88540 rattus norv
703	86	11	Q920Y5	10.9	776	5	10.9	101	16	Q9PPB2	Q9PPB2 campylobact
704	87	2	Q9AIF2	10.9	777	5	10.9	102	16	Q9KFB0	Q9KFB0 bacillus ha
705	87	5	Q93879	10.9	778	5	10.9	102	16	Q928K3	Q928K3 listeria in
706	87	10	Q94KC4	10.9	779	5	10.9	103	3	Q05457	Q05457 saccharomyc
707	87	16	Q97EY0	10.9	780	5	10.9	103	3	Q9NLT9	Q9NLT9 leishmania
708	88	2	Q9RQH4	10.9	781	5	10.9	103	5	O15941	O15941 bombyx mori
709	88	5	Q93880	10.9	782	5	10.9	103	10	Q9ZRW1	Q9ZRW1 cicor ariet
710	88	6	Q97588	10.9	783	5	10.9	103	12	Q9E8F7	Q9E8F7 porcine ade
711	88	10	Q9L199	10.9	784	5	10.9	103	15	Q71731	Q71731 human immun
712	89	2	Q93M16	10.9	785	5	10.9	103	16	Q9JSA8	Q9JSA8 chlamydia p
713	89	5	Q95S05	10.9	786	5	10.9	104	9	Q9AFC3	Q9AFC3 beta vulgar
714	89	13	Q90VX5	10.9	787	5	10.9	104	9	Q9AZH1	Q9AZH1 bacterioph
715	89	16	Q9HW24	10.9	788	5	10.9	104	10	Q93084	Q93084 arabisdopsis
716	89	16	Q98A25	10.9	789	5	10.9	104	16	Q9CEA3	Q9CEA3 lactococcus
717	90	5	Q9N124	10.9	790	5	10.9	104	16	Q992L7	Q992L7 streptococc
718	90	15	Q9JFL6	10.9	791	5	10.9	104	16	Q98111	Q98111 rhizobium l
719	90	16	Q98Q11	10.9	792	5	10.9	104	17	O57995	O57995 pyrococcus
720	90	17	Q977A4	10.9	793	5	10.9	105	12	Q98524	Q98524 paramecium
721	91	6	Q28857	10.9	794	5	10.9	105	16	Q9PQR7	Q9PQR7 ureaplasma
722	91	12	O02185	10.9	795	5	10.9	105	16	Q9JWB2	Q9JWB2 neisseria m
723	91	16	Q92LQ2	10.9	796	5	10.9	106	8	Q9MD41	Q9MD41 saccharomyc
724	92	3	Q01695	10.9	797	5	10.9	106	16	O50937	O50937 borrelia bu
725	92	5	P82116	10.9	798	5	10.9	106	16	Q9A169	Q9A169 streptococc
726	92	15	Q9WBE3	10.9	799	5	10.9	106	16	Q92G18	Q92G18 rickettsia
727	92	16	Q9K2M8	10.9	800	5	10.9	107	10	Q9M0B5	Q9M0B5 arabisdopsis
728	92	16	Q9JR94	10.9	801	5	10.9	107	11	Q9D5B6	Q9D5B6 mus musculu
729	93	2	Q55253	10.9	802	5	10.9	107	11	Q9D5B6	Q9D5B6 mus musculu
730	93	9	Q37978	10.9	803	5	10.9	107	15	Q75678	Q75678 human immun
731	93	16	Q52F49	10.9	804	5	10.9	107	15	Q75683	Q75683 human immun
732	94	2	Q56942	10.9	805	5	10.9	107	15	Q75684	Q75684 human immun
733	95	2	Q48582	10.9	806	5	10.9	107	17	O25702	O25702 archaeoglob
734	95	3	Q00953	10.9	807	5	10.9	107	17	Q73985	Q73985 pyrococcus
735	95	15	Q9W9W3	10.9	808	5	10.9	108	2	Q9F3G7	Q9F3G7 streptomyc
736	95	15	Q9W9N5	10.9	809	5	10.9	108	2	Q93417	Q93417 vibrio salm
737	95	15	Q73320	10.9	810	5	10.9	108	5	Q9VB37	Q9VB37 drosophila
738	95	17	Q28676	10.9	811	5	10.9	108	5	O02562	O02562 acrolepiops
739	96	2	Q9F6C4	10.9	812	5	10.9	108	6	Q9XST8	Q9XST8 canis fami
740	96	10	Q93VC1	10.9	813	5	10.9	108	10	Q9SXR1	Q9SXR1 oryza sativ
741	96	12	Q93180	10.9	814	5	10.9	108	10	Q940A0	Q940A0 arabisdopsis
742	96	12	Q9W131	10.9	815	5	10.9	108	16	O32892	O32892 mycobacteri
743	96	13	Q90X51	10.9	816	5	10.9	109	2	Q9L7X0	Q9L7X0 clostridium
744	97	10	Q42187	10.9	817	5	10.9	109	2	Q9ANF1	Q9ANF1 bradyrhizob
745	97	12	Q9E8G1	10.9	818	5	10.9	109	2	Q9X7N0	Q9X7N0 streptomyc
746	97	16	Q9K5P2	10.9	819	5	10.9	109	5	P91881	P91881 diadromus p

820	5	10.9	109	5	P91877	P91877 diprion pin	893	5	10.9	120	2	Q9ALG7	Q9ALG7 uncultured
821	5	10.9	109	5	Q9FV31	Q9FV31 pisum sativ	894	5	10.9	120	2	Q9ALG4	Q9ALG4 uncultured
822	5	10.9	110	9	Q38474	Q38474 bacterioph	895	5	10.9	120	2	Q9ALG1	Q9ALG1 uncultured
823	5	10.9	110	11	Q9D3T2	Q9D3T2 mus musculu	896	5	10.9	120	2	Q9ALG0	Q9ALG0 uncultured
824	5	10.9	110	11	Q61977	Q61977 mus musculu	897	5	10.9	120	2	Q9ALF9	Q9ALF9 uncultured
825	5	10.9	110	16	Q931X8	Q931X8 staphylococ	898	5	10.9	120	2	Q9ALF7	Q9ALF7 uncultured
826	5	10.9	111	5	Q9N126	Q9N126 plasmodium	899	5	10.9	120	2	Q9ALF5	Q9ALF5 uncultured
827	5	10.9	111	11	Q9D9Y7	Q9D9Y7 mus musculu	900	5	10.9	120	2	Q9ALF8	Q9ALF8 uncultured
828	5	10.9	111	16	Q9Z3A0	Q9Z3A0 escherichia	901	5	10.9	120	2	Q9ALF6	Q9ALF6 uncultured
829	5	10.9	112	5	Q9XVH1	Q9XVH1 caenorhabdi	902	5	10.9	120	2	Q9ALE8	Q9ALE8 uncultured
830	5	10.9	113	3	Q00745	Q00745 emericella	903	5	10.9	120	2	Q9ALE5	Q9ALE5 uncultured
831	5	10.9	113	12	Q89271	Q89271 carnation m	904	5	10.9	120	2	Q9ALE3	Q9ALE3 uncultured
832	5	10.9	113	13	Q91153	Q91153 notophthalm	905	5	10.9	120	2	Q9ALE0	Q9ALE0 uncultured
833	5	10.9	113	16	Q9KFU4	Q9KFU4 bacillus ha	906	5	10.9	120	2	Q9ALD9	Q9ALD9 uncultured
834	5	10.9	114	10	Q41211	Q41211 striga asia	907	5	10.9	120	2	Q9ALD8	Q9ALD8 uncultured
835	5	10.9	114	17	Q58083	Q58083 pyrococcus	908	5	10.9	120	2	Q9ALD7	Q9ALD7 uncultured
836	5	10.9	115	3	Q59983	Q59983 aspergillus	909	5	10.9	120	2	Q9ALD6	Q9ALD6 uncultured
837	5	10.9	115	3	Q59988	Q59988 aspergillus	910	5	10.9	120	2	Q9ALD5	Q9ALD5 uncultured
838	5	10.9	115	3	Q59989	Q59989 aspergillus	911	5	10.9	120	2	Q9ALD4	Q9ALD4 uncultured
839	5	10.9	115	3	Q59990	Q59990 neosartoria	912	5	10.9	120	2	Q9ALD3	Q9ALD3 uncultured
840	5	10.9	115	3	Q59991	Q59991 neosartoria	913	5	10.9	120	2	Q9ALD2	Q9ALD2 uncultured
841	5	10.9	115	3	Q59994	Q59994 neosartoria	914	5	10.9	120	2	Q9ALD1	Q9ALD1 uncultured
842	5	10.9	115	3	Q59995	Q59995 neosartoria	915	5	10.9	120	2	Q9ALD6	Q9ALD6 uncultured
843	5	10.9	115	3	Q59997	Q59997 neosartoria	916	5	10.9	120	2	Q9ALC1	Q9ALC1 uncultured
844	5	10.9	115	3	Q60001	Q60001 neosartoria	917	5	10.9	120	2	Q9ALC0	Q9ALC0 uncultured
845	5	10.9	115	3	Q60002	Q60002 neosartoria	918	5	10.9	120	2	Q9ALB9	Q9ALB9 uncultured
846	5	10.9	115	3	Q19616	Q19616 caenorhabdi	919	5	10.9	120	2	Q9ALB8	Q9ALB8 uncultured
847	5	10.9	115	9	Q64164	Q64164 bacterioph	920	5	10.9	120	2	Q9ALB7	Q9ALB7 uncultured
848	5	10.9	115	12	Q91161	Q91161 regina rana	921	5	10.9	120	2	Q9ALB1	Q9ALB1 uncultured
849	5	10.9	115	16	Q31884	Q31884 bacillus su	922	5	10.9	120	2	Q93TH2	Q93TH2 uncultured
850	5	10.9	115	16	Q9CFJ9	Q9CFJ9 lactococcus	923	5	10.9	120	2	Q93L61	Q93L61 uncultured
851	5	10.9	116	2	Q9EJH8	Q9EJH8 salmone	924	5	10.9	120	2	Q93KN6	Q93KN6 uncultured
852	5	10.9	116	5	Q9BJQ2	Q9BJQ2 plasmodium	925	5	10.9	120	2	Q9NAS2	Q9NAS2 brugia mala
853	5	10.9	116	8	Q9B8A4	Q9B8A4 trichinella	926	5	10.9	120	17	Q58311	Q58311 pyrococcus
854	5	10.9	116	13	Q9W787	Q9W787 ciemophorus	927	5	10.9	120	17	Q58601	Q58601 pyrococcus
855	5	10.9	116	16	Q9X1V2	Q9X1V2 thermotoga	928	5	10.9	121	2	Q9RNU0	Q9RNU0 solar lake
856	5	10.9	116	17	Q27245	Q27245 methanother	929	5	10.9	121	2	Q9RNT6	Q9RNT6 solar lake
857	5	10.9	117	5	Q9B4P9	Q9B4P9 plasmodium	930	5	10.9	121	2	Q9RNT4	Q9RNT4 solar lake
858	5	10.9	117	10	Q9SXR0	Q9SXR0 oryza sativ	931	5	10.9	121	2	Q9RNT3	Q9RNT3 solar lake
859	5	10.9	117	10	Q9S8S6	Q9S8S6 arabidopsis	932	5	10.9	121	2	Q9RNT1	Q9RNT1 solar lake
860	5	10.9	117	17	Q9XC28	Q9XC28 aeropyrum p	933	5	10.9	121	2	Q9RNS9	Q9RNS9 solar lake
861	5	10.9	118	2	Q9RNT8	Q9RNT8 solar lake	934	5	10.9	121	2	Q9RNS7	Q9RNS7 solar lake
862	5	10.9	118	10	Q9RTO7	Q9RTO7 arabidopsis	935	5	10.9	121	2	Q9RNS5	Q9RNS5 solar lake
863	5	10.9	118	10	Q9FYI2	Q9FYI2 arabidopsis	936	5	10.9	121	2	Q9RNS4	Q9RNS4 solar lake
864	5	10.9	118	10	Q949R3	Q949R3 arabidopsis	937	5	10.9	121	2	Q9RNS2	Q9RNS2 solar lake
865	5	10.9	118	12	Q9YW07	Q9YW07 melanoplus	938	5	10.9	121	2	Q9RNS1	Q9RNS1 solar lake
866	5	10.9	118	12	Q9YMG9	Q9YMG9 lymantria d	939	5	10.9	121	2	Q9RMR3	Q9RMR3 solar lake
867	5	10.9	118	16	Q988K6	Q988K6 rhizobium l	940	5	10.9	121	2	Q9RMR1	Q9RMR1 solar lake
868	5	10.9	118	16	Q9AKJ5	Q9AKJ5 rhizobium m	941	5	10.9	121	2	Q9ALH8	Q9ALH8 uncultured
869	5	10.9	119	2	Q9XD86	Q9XD86 corynebacte	942	5	10.9	121	2	Q9ALH5	Q9ALH5 uncultured
870	5	10.9	119	2	Q9ALB4	Q9ALB4 uncultured	943	5	10.9	121	2	Q9ALG6	Q9ALG6 uncultured
871	5	10.9	119	2	Q9ALE2	Q9ALE2 uncultured	944	5	10.9	121	2	Q9ALF8	Q9ALF8 uncultured
872	5	10.9	119	2	Q9ALE1	Q9ALE1 uncultured	945	5	10.9	121	2	Q9ALF3	Q9ALF3 uncultured
873	5	10.9	119	2	Q9ALC9	Q9ALC9 uncultured	946	5	10.9	121	2	Q9ALF1	Q9ALF1 uncultured
874	5	10.9	119	2	Q9ALC8	Q9ALC8 uncultured	947	5	10.9	121	2	Q9ALF0	Q9ALF0 uncultured
875	5	10.9	119	5	Q9GQZ0	Q9GQZ0 plasmodium	948	5	10.9	121	2	Q9ALD0	Q9ALD0 uncultured
876	5	10.9	119	10	Q9TOL7	Q9TOL7 arabidopsis	949	5	10.9	121	2	Q9ALC5	Q9ALC5 uncultured
877	5	10.9	119	16	Q67530	Q67530 aquifex aeo	950	5	10.9	121	2	Q9ALC4	Q9ALC4 uncultured
878	5	10.9	119	17	Q980M9	Q980M9 sulfobobus	951	5	10.9	121	2	Q9ALC2	Q9ALC2 uncultured
879	5	10.9	120	2	Q9RNT9	Q9RNT9 solar lake	952	5	10.9	121	2	Q9ALB5	Q9ALB5 uncultured
880	5	10.9	120	2	Q9RNT5	Q9RNT5 solar lake	953	5	10.9	121	2	Q9ALB4	Q9ALB4 uncultured
881	5	10.9	120	2	Q9RNT2	Q9RNT2 solar lake	954	5	10.9	121	2	Q9ALB3	Q9ALB3 uncultured
882	5	10.9	120	2	Q9RNT0	Q9RNT0 solar lake	955	5	10.9	121	2	Q9ALB2	Q9ALB2 uncultured
883	5	10.9	120	2	Q9ALH9	Q9ALH9 uncultured	956	5	10.9	121	5	P91789	P91789 taenia ovis
884	5	10.9	120	2	Q9ALH7	Q9ALH7 uncultured	957	5	10.9	121	5	Q95YV8	Q95YV8 leishmania
885	5	10.9	120	2	Q9ALH6	Q9ALH6 uncultured	958	5	10.9	121	12	Q41658	Q41658 influenza a
886	5	10.9	120	2	Q9ALH4	Q9ALH4 uncultured	959	5	10.9	121	15	Q9WNY0	Q9WNY0 human immun
887	5	10.9	120	2	Q9ALH3	Q9ALH3 uncultured	960	5	10.9	122	2	Q9RNT7	Q9RNT7 solar lake
888	5	10.9	120	2	Q9ALH2	Q9ALH2 uncultured	961	5	10.9	122	2	Q9RNS3	Q9RNS3 solar lake
889	5	10.9	120	2	Q9ALH1	Q9ALH1 uncultured	962	5	10.9	122	5	Q9TYG7	Q9TYG7 plectus acu
890	5	10.9	120	2	Q9ALH0	Q9ALH0 uncultured	963	5	10.9	122	16	Q9JTV4	Q9JTV4 neisseria m
891	5	10.9	120	2	Q9ALG9	Q9ALG9 uncultured	964	5	10.9	123	2	Q9ALB6	Q9ALB6 uncultured
892	5	10.9	120	2	Q9ALG8	Q9ALG8 uncultured	965	5	10.9	123	5	Q9V8Q4	Q9V8Q4 drosophila

RESULT 4
Q9R7H6 PRELIMINARY; PRT; 363 AA.
AC Q9R7H6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DE HYPOTHETICAL 41.9 KDA PROTEIN (FRAGMENT)
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N40;
RA Feng S., Chen G., Barthold S.;
RT "G4B clone from Borrelia burgdorferi."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF005055; AAD01225.1; -;
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 363 AA; 41909 MW; D8330154940932CA CRC64;

Query Match 19.6%; Score 9; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TSLKNDKF 12
DB 45 TSLKNDKF 53

RESULT 5
Q98RG2 PRELIMINARY; PRT; 88 AA.
AC Q98RG2;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DE HYPOTHETICAL PROTEIN MYPV_0470.
GN MYPV_0470.
OS Mycoplasma pulmonis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2107;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UAB CTIP;
RX MEDLINE=21267165; PubMed=11353084;
RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,
RA Moszer I., Dybvig K., Wroblewski H., Vlari A., Rocha E.P.C.,
RA Blanchard A.;
RT "The complete genome sequence of the murine respiratory pathogen
RT Mycoplasma pulmonis";
RL Nucleic Acids Res. 29:2145-2153(2001).
DR EMBL; AL445563; CAC13220.1; -;
DR Mypulist; MYPV_0470; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 88 AA; 10502 MW; 49ECE9BD060A3407 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 88;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KTLKLN 9
DB 20 KTLKLN 26

RESULT 6
Q83163 PRELIMINARY; PRT; 96 AA.
ID Q83163

Q83163;
AC 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE (ALTERED VIRULENCE ISOLATE D/H), COMPLETE GENOME.
OS Cauliflower mosaic virus.
OC Viruses; Retroid viruses; Caulimovirus.
OX NCBI_TaxID=10641;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CABB-D/H;
RX MEDLINE=83106468; PubMed=7152260;
RA Balazs E., Guilly H., Jonard G., Richards K.;
RT "Nucleotide sequence of DNA from an altered-virulence isolate D/H of
RT the cauliflower mosaic virus."
RL Gene 19:239-249(1982).
DR EMBL; M10376; AAA46344.1; -;
SQ SEQUENCE 96 AA; 11429 MW; F43565C5790A7657 CRC64;

Query Match 15.2%; Score 7; DB 12; Length 96;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KTLKLN 9
DB 75 KTLKLN 81

RESULT 7
Q96Y40 PRELIMINARY; PRT; 239 AA.
AC Q96Y40;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE HYPOTHETICAL PROTEIN ST2326.
GN ST2326
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo H. I., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain 7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000989; BAB67437.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 239 AA; 28001 MW; 1718022FEAD067EC CRC64;

Query Match 15.2%; Score 7; DB 17; Length 239;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 SLKNDKF 12
DB 148 SLKNDKF 154

RESULT 8
Q9I488 PRELIMINARY; PRT; 240 AA.
ID Q9I488;
AC Q9I488;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)


```

DE PROBABLE ATP-BINDING COMPONENT OF ABC TRANSPORTER.
GN PA1256.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Gollity L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:950-964(2000);
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
CC (ABC TRANSPORTERS).
DR EMBL; AE004555; AAG04645.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_Grp_A.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Complete proteome; Transport.
SQ SEQUENCE 240 AA; 26548 MW; 3F6796D7C10F0F83 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 240;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
Db 32 IIGSGS 38
|||||
RESULT 9
ID Q92W53 PRELIMINARY; PRT; 240 AA.
AC Q92W53;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE AMINO ACID ABC TRANSPORTER ATP-BINDING PROTEIN.
GN SMB20266.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL603642; CAC48656.1; -.
DR ATP-binding; Plasmid; Hypothetical protein; Complete proteome.
KW ATP-binding; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 240 AA; 26366 MW; E9826924BC60E597 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 240;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
Db 32 IIGSGS 38
|||||
RESULT 10
ID Q90315 PRELIMINARY; PRT; 253 AA.
AC Q90315;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE STEM CELL FACTOR, MEMBRANE-BOUND FORM.
OS Coturnix coturnix (Common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96283808; PubMed=8679698;
RA Pettitte J.N., Kulik M.J.;
RT "Cloning and characterization of cDNAs encoding two forms of avian
RT stem cell factor.";
RL Biochim. Biophys. Acta 1307:149-151(1996).
DR EMBL; U43079; AAC59934.1; -.
DR InterPro: IPR003452; SCF.
DR Pfam: PF02404; SCF; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Complete proteome; Transport.
SQ SEQUENCE 253 AA; 28993 MW; 20709B3854F0207A CRC64;

Query Match 15.2%; Score 7; DB 13; Length 253;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 FFRLFNR 38
Db 144 FFRLFNR 150
|||||
RESULT 11
ID O67894 PRELIMINARY; PRT; 253 AA.
AC O67894;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ABC TRANSPORTER.
GN ABC12 OR AQ_2137.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=5537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Auway M., Auway R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL; AE000774; AAC07853.1; -.
DR HSSP; Q58663; IG6H.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR Pfam: PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00382; AAA; 1.
KW Complete proteome.
SQ SEQUENCE 253 AA; 28621 MW; C880AE769CF8A4B2 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 253;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
Db 32 IIGSGS 38
|||||

```



```

QY 22 IIGSGS 28
    |||||
Db 35 IIGSGS 41

RESULT 12
Q98CC9 PRELIMINARY; PRT; 268 AA.
AC Q98CC9;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
DE AMINO ACID ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN MLL5204.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003006; BAB51692.1; -.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transport.
DR InterPro; IPR001687; ATP_GTP_A.
DR Pfam; PF00005; ABC_tran; 1.
DR SMART; SM00382; AAA; 1.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 268 AA; 29407 MW; 8676DDE8ED78FD34 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 268;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
    |||||
Db 49 IIGSGS 55

RESULT 13
Q9SEW3 PRELIMINARY; PRT; 276 AA.
AC Q9SEW3;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE RECEPTOR-LIKE PROTEIN KINASE (FRAGMENT).
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WILLIAMS 82;
RA Bassuner R., Cao H., Mackenzie S.;
RT "A novel receptor-like protein kinase from soybean and its relation to
RT the resistance phenotype to cyst nematode infection."
RL Submitted (MAY1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF068135; AAF21775.1; -.
DR HSSP; P05046; 1SBE.
DR InterPro; IPR000985; Lectin_lega.
DR Pfam; PF00138; lectin_lega; 1.

```

```

KW Kinase.
FT NON_TER 276
FT NON_TER 276
SQ SEQUENCE 276 AA; 30521 MW; FAFB0769E0D1114 CRC64;

Query Match 15.2%; Score 7; DB 10; Length 276;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KTLCLKN 9
    |||||
Db 226 KTLCLKN 232

RESULT 14
Q92S62 PRELIMINARY; PRT; 284 AA.
AC Q92S62;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE PUTATIVE ATP-BINDING ABC TRANSPORTER PROTEIN.
GN SMC02260.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21368234; PubMed=11474104;
RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,
RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,
RA Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P.,
RA Cloux A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,
RA Gloux S., Godrie T., Goffeau A., Golding B., Guzy J., Gurjal M.,
RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,
RA Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure V.,
RA Masny D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,
RA Ramseger U., Surzycki R., Thebault P., Vandenbol M.,
RA Vorhoefer F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;
RT "The composite genome of the legume symbiont Sinorhizobium meliloti."
RL Science 293:668-672(2001).
DR EMBL; AL591784; CAC45136.1; -.
KW Complete proteome.
SQ SEQUENCE 284 AA; 31218 MW; 82E1AFB4DAA493C1 CRC64;

Query Match 15.2%; Score 7; DB 16; Length 284;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 IIGSGS 28
    |||||
Db 65 IIGSGS 71

RESULT 15
Q90314 PRELIMINARY; PRT; 287 AA.
AC Q90314;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE STEM CELL FACTOR.
OS Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9628308; PubMed=8679698;
RA Pettite J.N., Kulik M.J.;

```

RT "Cloning and characterization of cDNAs encoding two forms of avian

stem cell factor.",

RL Biochim. Biophys. Acta 1307:149-151(1996).

DR EMBL: U43078; AAC59933.1; -.

DR InterPro: IPR003452; SCF.

DR Pfam: PF02404; SCF; 1.

SQ SEQUENCE 287 AA; 32455 MW; ABA81AEM422A702E CRC64;

Query Match

15.2%; Score 7; DB 13; Length 287;

Best Local Similarity 100.0%; Pred. No. 29;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 FFRLFN 38

|||||

Db 144 FFRLFN 150

Search completed: November 5, 2002, 11:06:28

Job time : 87.209 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:02 ; Search time 16.9254 seconds
(without alignments)
137.814 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 104

Sequence: 1 SGLSTFRLFRNSTQALGK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	46	44.2	577	21	Arabidopsis thalia
2	46	44.2	596	21	Arabidopsis thalia
3	46	44.2	660	21	Arabidopsis thalia
4	45	43.3	391	22	Arabidopsis thalia
5	44	42.3	340	22	Murine 3-OST-3B.
6	44	42.3	340	22	Human gene 18 enco
7	44	42.3	340	22	Renal and cardiova
8	44	42.3	340	22	Renal and cardiova
9	43	41.3	86	22	Human protein sequ
10	43	41.3	106	19	Human polypeptide
11	43	41.3	135	14	DNA encoding Staph
					Serotonin receptor

12	43	41.3	151	19	AAW79358	DNA encoding Staph
13	43	41.3	167	22	AAU34248	Staphylococcus aur
14	43	41.3	167	22	AAU36819	Staphylococcus aur
15	43	41.3	167	22	AAU37429	Staphylococcus aur
16	43	41.3	479	14	AAK41942	Serotonin receptor
17	43	41.3	504	15	AAK54682	Mouse brain 5HT2C
18	42	40.4	43	22	AAO06271	Human polypeptide
19	42	40.4	114	22	AAK93004	Human protein sequ
20	42	40.4	148	22	AAK93004	Human protein sequ
21	42	40.4	225	22	ABB65926	Drosophila melanog
22	42	40.4	443	21	AAK42806	Human OREF ORF2570
23	42	40.4	453	21	AAK68799	Amino acid sequenc
24	42	40.4	482	22	AAK51776	Novel human secret
25	42	40.4	499	22	ABK62187	Drosophila melanog
26	41	39.4	148	22	ABG28821	Novel human diagno
27	41	39.4	183	18	AAW28302	Amino acid sequenc
28	41	39.4	250	21	AAK58800	Breast and ovarian
29	41	39.4	437	19	AAW53264	Human adipocyte-sp
30	41	39.4	437	21	AAK99534	Human adipocyte-sp
31	41	39.4	595	20	AAK27236	Amino acid sequenc
32	41	39.4	619	20	AAK27236	Amino acid sequenc
33	41	39.4	761	20	AAK27211	Amino acid sequenc
34	41	39.4	1798	22	ABG09933	Novel human diagno
35	41	39.4	1981	19	AAW42634	Protein sequence t
36	41	39.4	2001	22	AAK72649	Human ATP binding
37	41	39.4	2436	22	AAK67160	Amino acid sequenc
38	41	39.4	2436	22	AAK62210	Human ARCA2 transp
39	40	38.5	48	22	AAK76591	Human colon cancer
40	40	38.5	52	21	AAK43141	Human OREF ORF2905
41	40	38.5	70	22	AAU02768	Mutant SCF polypep
42	40	38.5	70	22	AAU02768	Mutant SCF polypep
43	40	38.5	82	22	AAU05264	Mutant SCF polypep
44	40	38.5	82	22	AAK98366	Human SCF related
45	40	38.5	82	22	AAU02459	Mutant SCF polypep

ALIGNMENTS

RESULT 1
AAG48997
ID AAG48997 standard; Protein: 577 AA.
AC AAG48997;
XX
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID: NO: 61937.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
FN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.

PR	23-APR-1999;	990S-0130510.	990S-0144632.	PR	20-JUL-1999;	990S-0144632.
PR	23-APR-1999;	990S-0130891.	990S-0144864.	PR	20-JUL-1999;	990S-0144864.
PR	28-APR-1999;	990S-0131449.	990S-0145086.	PR	21-JUL-1999;	990S-0145086.
PR	30-APR-1999;	990S-0132048.	990S-0145088.	PR	21-JUL-1999;	990S-0145088.
PR	04-MAY-1999;	990S-0132484.	990S-0145085.	PR	22-JUL-1999;	990S-0145085.
PR	05-MAY-1999;	990S-0132485.	990S-0145087.	PR	22-JUL-1999;	990S-0145087.
PR	06-MAY-1999;	990S-0132486.	990S-0145089.	PR	22-JUL-1999;	990S-0145089.
PR	07-MAY-1999;	990S-0132487.	990S-0145192.	PR	22-JUL-1999;	990S-0145192.
PR	11-MAY-1999;	990S-0132863.	990S-0145145.	PR	23-JUL-1999;	990S-0145145.
PR	14-MAY-1999;	990S-0134256.	990S-0145224.	PR	23-JUL-1999;	990S-0145224.
PR	14-MAY-1999;	990S-0134218.	990S-0145226.	PR	26-JUL-1999;	990S-0145217.
PR	14-MAY-1999;	990S-0134219.	990S-0145213.	PR	27-JUL-1999;	990S-0145213.
PR	18-MAY-1999;	990S-0134370.	990S-0145918.	PR	27-JUL-1999;	990S-0145918.
PR	18-MAY-1999;	990S-0134376.	990S-0145919.	PR	27-JUL-1999;	990S-0145919.
PR	19-MAY-1999;	990S-0134941.	990S-0145951.	PR	28-JUL-1999;	990S-0145951.
PR	20-MAY-1999;	990S-0135124.	990S-0146386.	PR	02-AUG-1999;	990S-0146386.
PR	21-MAY-1999;	990S-0135353.	990S-0146388.	PR	02-AUG-1999;	990S-0146388.
PR	24-MAY-1999;	990S-0135629.	990S-0146389.	PR	02-AUG-1999;	990S-0146389.
PR	25-MAY-1999;	990S-0136021.	990S-0147038.	PR	03-AUG-1999;	990S-0147038.
PR	27-MAY-1999;	990S-0136392.	990S-0147204.	PR	04-AUG-1999;	990S-0147204.
PR	28-MAY-1999;	990S-0136782.	990S-0147302.	PR	04-AUG-1999;	990S-0147302.
PR	01-JUN-1999;	990S-0137222.	990S-0147192.	PR	05-AUG-1999;	990S-0147192.
PR	03-JUN-1999;	990S-0137528.	990S-0147303.	PR	05-AUG-1999;	990S-0147260.
PR	04-JUN-1999;	990S-0137502.	990S-0147303.	PR	06-AUG-1999;	990S-0147303.
PR	07-JUN-1999;	990S-0137724.	990S-0147416.	PR	06-AUG-1999;	990S-0147416.
PR	08-JUN-1999;	990S-0138094.	990S-0147493.	PR	06-AUG-1999;	990S-0147493.
PR	10-JUN-1999;	990S-0138540.	990S-0147935.	PR	09-AUG-1999;	990S-0147935.
PR	10-JUN-1999;	990S-0138847.	990S-0148171.	PR	10-AUG-1999;	990S-0148171.
PR	14-JUN-1999;	990S-0139119.	990S-0148319.	PR	11-AUG-1999;	990S-0148319.
PR	16-JUN-1999;	990S-0139452.	990S-0148341.	PR	12-AUG-1999;	990S-0148341.
PR	16-JUN-1999;	990S-0139453.	990S-0148565.	PR	13-AUG-1999;	990S-0148565.
PR	17-JUN-1999;	990S-0139492.	990S-0148684.	PR	13-AUG-1999;	990S-0148684.
PR	18-JUN-1999;	990S-0139454.	990S-0149368.	PR	16-AUG-1999;	990S-0149368.
PR	18-JUN-1999;	990S-0139455.	990S-0149172.	PR	17-AUG-1999;	990S-0149172.
PR	18-JUN-1999;	990S-0139456.	990S-0149426.	PR	17-AUG-1999;	990S-0149426.
PR	18-JUN-1999;	990S-0139457.	990S-0149722.	PR	18-AUG-1999;	990S-0149722.
PR	18-JUN-1999;	990S-0139458.	990S-0149723.	PR	20-AUG-1999;	990S-0149723.
PR	18-JUN-1999;	990S-0139459.	990S-0149929.	PR	20-AUG-1999;	990S-0149929.
PR	18-JUN-1999;	990S-0139460.	990S-0149930.	PR	23-AUG-1999;	990S-0149902.
PR	18-JUN-1999;	990S-0139461.	990S-0150566.	PR	23-AUG-1999;	990S-0149930.
PR	18-JUN-1999;	990S-0139462.	990S-0150884.	PR	25-AUG-1999;	990S-0150566.
PR	18-JUN-1999;	990S-0139463.	990S-0150884.	PR	26-AUG-1999;	990S-0150884.
PR	18-JUN-1999;	990S-0139750.	990S-0151065.	PR	27-AUG-1999;	990S-0151065.
PR						

PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.
 PR 29-OCT-1999; 99US-0162143.

Query Match 44.28; Score 46; DB 21; Length 577;

Best Local Similarity 69.28; Pred. No. 24;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 FFRLENRSTQAL 19

Db 411 FFRLEDRDRQRL 423

RESULT 2

ID AAG48996
 AC AAG48996; Protein: 596 AA.

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 61936.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126284.

PR 01-APR-1999; 99US-0126785.

PR 06-APR-1999; 99US-0127452.

PR 08-APR-1999; 99US-0128234.

PR 16-APR-1999; 99US-0128714.

PR 19-APR-1999; 99US-0129845.

PR 21-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 18-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.

```
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

Query Match 44.2%; Score 46; DB 21; Length 596;
Best Local Similarity 69.2%; Pred. No. 25;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 FFRLENRSTQAL 19
Db 430 FERLEDRFRQL 442

RESULT 3
AAG48995
ID AAG48995 standard; Protein; 660 AA.
XX
AC AAG48995;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61935.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW Hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
```


PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161360.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 44.2%; Score 46; DB 21; Length 660;

Best Local Similarity 69.2%; Pred. No. 28;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 FFLRNSETQAL 19

|||||:| | |

Db 494 FFLRDRFRQL 506

RESULT 4

AAB70114

ID AAB70114 standard; protein; 391 AA.

XX AAB70114;

XX AAB70114;

DT 17-MAY-2001 (first entry)

XX Murine 3-OST-3B.

DE Mouse; 3-OST-3B; heparan sulphate 3-O-sulphotransferase 3B; antiviral;

KW herpes simplex virus type-1; HSV-1; infection.

KW Mus sp.

OS WO200113910-A2.

PN 01-MAR-2001.

XX 25-AUG-2000; 2000WO-US23288.

XX 25-AUG-1999; 99US-0150743.

PR 10-JAN-2000; 2000US-0175347.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

PA (NOUN) UNIV NORTHWESTERN.

XX Shukla D, Liu J, Rosenberg RD, Spear PG;

XX WPI; 2001-234971/24.

DR Novel pharmaceutical preparation useful for inhibiting herpes simplex

XX virus type-1 infection in a mammal, comprises a polysaccharide

PT preparation enriched for 3-O-sulfated glycosamine residues

PT Disclosure; Fig 1; 29pp; English.

XX The present sequence is murine heparan sulphate 3-O-sulphotransferase

CC (3-OST)-3B. A substantially pure polysaccharide preparation enriched for

CC 3-OST-3 modified glucosamine residues is useful for inhibiting herpes

CC simplex virus type-1 (HSV-1) infection in a mammal. It is useful for

CC inhibiting viral penetration and entry of HSV-1 into mammalian cells

CC through heparan sulphate binding.

XX Sequence 391 AA;

SQ

Query Match 43.3%; Score 45; DB 22; Length 391;

Best Local Similarity 50.0%; Pred. No. 23;

Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 LSTFFRLNRSFTQALGK 21

| | | | | | | |

Db 369 LRDFYRPNKRYQMTGR 386

RESULT 5
 AAE09656
 ID AAE09656 standard; protein; 340 AA.
 XX
 AC AAE09656;
 XX
 DT 19-NOV-2001 (first entry)
 XX
 DE Human gene 18 encoded lipid metabolism protein HDACA29, SEQ ID NO:50.
 XX
 KW Human; lipid metabolism protein; nontropic; neuroprotective; cardiant;
 KW cerebroprotective; hepatotropic; antidiabetic; ophthalmic; nephrotropic;
 KW immune disorder; autoimmune disease; rheumatoid arthritis; glossitis;
 KW systemic lupus erythematosus; conjunctivitis; inflammatory disorder;
 KW respiratory disorder; asthma; allergy; CNS disorder; Alzheimer's disease;
 KW Parkinson's disease; atherosclerosis; cardiovascular disorder; cancer;
 KW coronary disease; familial hypercholesterolaemia; hyperlipidaemia;
 KW haematopoietic disorder; hypolipidaemia; lipodosis; Gaucher's disease;
 KW Tay-sach's disease; mental retardation; gene therapy; antisense therapy.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 161 /label= Unknown
 FT /note= "Encoded by MAA"
 FT Misc-difference 278 /label= Unknown
 FT /note= "Encoded by AAN"
 FT Misc-difference 293 /label= Unknown
 FT /note= "Encoded by GAW"
 XX
 PN WO200155203-A1.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01327.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.

PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-457716/49.
XX N-PSDB; AAD16623.
DR Isolated lipid metabolism polypeptide for screening to identify
XX antagonists and agonists that may enhance or block activities mediated
XX by lipid metabolism proteins and also for testing and detection e.g.
XX diagnosis -
XX
XX Claim 11; SEQ ID No 50; 453pp; English.
XX
XX AAD16606-AAD16627 represent cDNAs corresponding to human lipid metabolism
XX protein genes, and AAE09639-AAE09660 represent the proteins they encode.
XX AAD16628-AAD16647 represent human lipid metabolism genomic DNAs. The
XX genes and their corresponding proteins are useful for preventing,
XX treating or ameliorating medical conditions such as immune disorders,
XX inflammatory disorders (e.g., conjunctivitis, glossitis), autoimmune
XX diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus),
XX respiratory disorders (asthma, allergy), CNS disorders (e.g., Alzheimer's
XX disease, Parkinson's disease), haematopoietic disorders, cardiovascular
XX disorders (e.g., atherosclerosis, coronary disease), hyperlipidaemias,
XX hypolipidaemia, lipidoses, familial hypercholesterolaemia, Gaucher's
XX disease, Tay-sach's disease, diabetes mellitus, cancer and other enzyme
XX
Query Match 42.3%; Score 44; DB 22; Length 340;
Best Local Similarity 42.9%; Pred. No. 30;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 SGLSTFFRLFNRSFTQALGK 21
|||:::
Db 40 SGYCSYFRAGSKPFPVLGE 60
RESULT 6
AAU18677
ID AAU18677 standard; Protein; 340 AA.
XX
AC AAU18677;
XX
XX 21-NOV-2001 (first entry)
DT


```
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0248477.
PR 08-NOV-2000; 2000US-0248478.
PR 08-NOV-2000; 2000US-0248523.
PR 08-NOV-2000; 2000US-0248524.
PR 08-NOV-2000; 2000US-0248525.
PR 08-NOV-2000; 2000US-0248526.
PR 08-NOV-2000; 2000US-0248527.
PR 08-NOV-2000; 2000US-0248528.
PR 08-NOV-2000; 2000US-0248532.
PR 08-NOV-2000; 2000US-0248533.
PR 08-NOV-2000; 2000US-0248609.
PR 08-NOV-2000; 2000US-0248610.
PR 08-NOV-2000; 2000US-0248611.
PR 08-NOV-2000; 2000US-0248613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 01-DEC-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.
XX ROSEN CA, Barash SC, Ruben SM;
XX WPI; 2001-488787/53.
XX N-PSDB; AAS30233.
XX New polynucleotides and polypeptides, useful for diagnosing, treating,
XX preventing or prognosing e.g. kidney, cardiovascular, blood,
XX electrolyte imbalance or neoplastic disorders, autoimmune diseases,
XX cancers
XX Claim 1; SEQ ID No 151; 506pp; English.
XX The invention relates to novel nucleic acids and polypeptides useful for
XX diagnosing, treating, preventing and/or prognosing disorders related to
XX these polypeptides. The polynucleotides are especially useful in the
XX diagnosis, prognosis, prevention and/or treatment of diseases which
XX include kidney disorders (e.g. renal failure or nephritis),
XX cardiovascular disorders (e.g. hypertension or myocardial infarction),
XX blood disorders (e.g. anaemia or blood coagulation disorders),
XX electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),
XX neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune
XX diseases, cancers, inflammatory diseases, reproductive system
XX disorders, endocrine disorders, neural activity and neurological
XX disorders, wound healing and respiratory disorders. AAU18644-AAU18715
XX represent the novel human renal and cardiovascular-associated amino
XX acid sequences of the invention. Note: The sequence data for this patent
XX did not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at:
XX ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 340 AA;
XX Query Match 42.3%; Score 44; DB 22; Length 340;
XX Best Local Similarity 42.9%; Pred. No. 30;
XX Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
QY 1 SGSLSTPFRLEFRSFTQALGK 21
DB 40 SGCYSTYFRAGSKPFNPVLGE 60
XX
RESULT 8
AAB95255
ID AAB95255 standard; Protein; 719 AA.
XX
AC AAB95255;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17425.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EPI074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
```


PS Claim 5; Pages 58-59; 114pp; English.

XX AAW79358-59 represent *Staphylococcus aureus* WCHU (NCIMB 40771)

CC proteins that have homology to a *Bacillus subtilis* single-strand

CC binding protein (helix-destabilising protein) (AAW79358) and a *Rattus*

CC norvegicus and a *Eubacterium ssDNA* binding protein (AAW79359). The

CC protein is used to generate antibodies and to screen for antimicrobials.

CC The products are used to treat or prevent bacterial infections,

CC particularly where caused by *S. aureus* but also against *Helicobacter*

CC *pylori*. Particular applications are to treat subjects before surgery or

CC insertion of an in-dwelling device (alternatively the device itself is

CC impregnated before placement). The nucleic acid sequence is used as

CC sources of antisense sequences (for therapeutic use) or regulatory

CC elements for controlling expression of bacterial genes, and for

CC antibacterial screening. The protein can be also used as a vaccine.

XX Sequence 106 AA;

SQ

Query Match 41.3%; Score 43; DB 19; Length 106;

Best Local Similarity 47.4%; Pred. No. 13;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 SLSTFFRLFNRSFTQALGK 21

DB 25 SVATFTLAVNRTFTNAQGE 43

II::II II::II I:

RESULT 11

AAW79358

ID AAR41945 standard; Protein; 135 AA.

XX AAR41945;

XX 10-MAY-1994 (first entry)

XX Serotonin receptor 5-HT2f C-terminal region.

XX

XX 5-hydroxytryptamine; receptor; serotoninergic; vasoconstriction;

KW plasmid pHD5HT2f; probe.

KW

XX EP565370-A.

XX

XX 13-OCT-1993.

XX

XX 07-APR-1993; 93EP-0302759.

XX

XX 09-APR-1992; 92US-0864005.

XX

XX (ELIL) LILLY & CO ELI.

XX

XX Baez M, Kursar JD;

XX

XX WPI; 1993-322574/41.

DR N-PSDB; AAQ49783.

DR

XX 5-Hydroxy-tryptamine receptor - used to identify drugs with

PT receptor activity

XX

XX Disclosure; Page 17-18; 20pp; English.

XX

XX The plasmid pHD5HT2f comprises the sequence AAQ49781 which codes for a

CC novel serotonin receptor. Fragments of the full-length coding

CC region are disclosed as suitable for use as probes to find

CC homologous (receptor) sequences. The preferred fragments are those

CC coding for the G-loop, the N-terminal and the C-terminal of the

CC 5-HT2f receptor (AAR41943-R41945, respectively).

XX

XX Sequence 135 AA;

SQ

Query Match 41.3%; Score 43; DB 14; Length 135;

Best Local Similarity 38.1%; Pred. No. 16;

Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 1 SGSLTFRLFNRSFTQALGK 21

DB 28 GVNPLIYTLFNKTFREAFGR 48

II::II II::II I:

RESULT 12

AAW79358

ID AAW79358 standard; Protein; 151 AA.

XX AAW79358;

XX 24-NOV-1998 (first entry)

XX DNA encoding *Staphylococcus aureus* single-strand binding protein.

DE

XX *Bacillus subtilis*; single-strand binding protein;

KW helix-destabilising protein; *Rattus norvegicus*; *Eubacterium*;

KW ssDNA binding protein; treatment; prevention; bacterial infection;

KW *Helicobacter pylori*; vaccine.

XX

OS *Staphylococcus aureus*.

XX

PN WO9823738-A2.

XX

PD 04-JUN-1998.

XX

PF 24-NOV-1997; 97WO-US22092.

XX

PR 25-NOV-1996; 96US-0031469.

XX

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX

XX Warren RL;

XX

XX WPI; 1998-322718/28.

DR N-PSDB; AAV59888.

XX

XX New nucleic acid from *Staphylococcus aureus* NCIMB 40771 - useful

PT for, e.g. diagnosis, prevention and treatment of bacterial

PT infection(s)

XX

XX Claim 5; Page 58; 114pp; English.

XX

XX AAW79358-59 represent *Staphylococcus aureus* WCHU (NCIMB 40771)

CC proteins that have homology to a *Bacillus subtilis* single-strand

CC binding protein (helix-destabilising protein) (AAW79358) and a *Rattus*

CC norvegicus and a *Eubacterium ssDNA* binding protein (AAW79359). The

CC protein is used to generate antibodies and to screen for antimicrobials.

CC The products are used to treat or prevent bacterial infections,

CC particularly where caused by *S. aureus* but also against *Helicobacter*

CC *pylori*. Particular applications are to treat subjects before surgery or

CC insertion of an in-dwelling device (alternatively the device itself is

CC impregnated before placement). The nucleic acid sequence is used as

CC sources of antisense sequences (for therapeutic use) or regulatory

CC elements for controlling expression of bacterial genes, and for

CC antibacterial screening. The protein can be also used as a vaccine.

XX

XX Sequence 151 AA;

SQ

Query Match 41.3%; Score 43; DB 19; Length 151;

Best Local Similarity 47.4%; Pred. No. 18;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 SLSTFFRLFNRSFTQALGK 21

DB 25 SVATFTLAVNRTFTNAQGE 43

II::II II::II I:

RESULT 13

AAU34248

ID AAU34248 standard; Protein; 167 AA.

XX

XX AAU34248;

XX DT 14-FEB-2002 (first entry)
 XX DE Staphylococcus aureus cellular proliferation protein #524.
 XX KW Antisense; prokaryotic cellular proliferation protein;
 XX KW antibiotic; antibacterial; drug design.
 XX OS Staphylococcus aureus.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US09180.
 XX PR 21-MAR-2000; 2000US-191078P.
 XX PR 23-MAY-2000; 2000US-206848P.
 XX PR 26-MAY-2000; 2000US-207727P.
 XX PR 23-OCT-2000; 2000US-242578P.
 XX PR 27-NOV-2000; 2000US-253625P.
 XX PR 22-DEC-2000; 2000US-257931P.
 XX PR 16-FEB-2001; 2001US-269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX DR WPI; 2001-611495/70.
 XX DR N-PSDB; AAS52107.
 XX PT New polynucleotides for the identification and development of
 XX PT antibiotics, comprise sequences of antisense nucleic acids -
 XX PS Example 3; Seq ID No 5744; 51lpp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 167 AA;

Query Match 41.3%; Score 43; DB 22; Length 167;
 Best Local Similarity 47.4%; Pred. No. 21;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 SLSTFFFLFNRSFTQALGK 21
 |::||| ||||| |:
 Db 25 SVATFTLVNRTFTNAQGE 43

RESULT 14
 AAU36819
 ID AAU36819 standard; Protein; 167 AA.
 XX
 AC AAU36819;

XX DT 14-FEB-2002 (first entry)
 XX DE Staphylococcus aureus cellular proliferation protein #989.
 XX KW Antisense; prokaryotic cellular proliferation protein;
 XX KW antibiotic; antibacterial; drug design.
 XX OS Staphylococcus aureus.
 XX PN WO200170955-A2.
 XX PD 27-SEP-2001.
 XX PF 21-MAR-2001; 2001WO-US09180.
 XX PR 21-MAR-2000; 2000US-191078P.
 XX PR 23-MAY-2000; 2000US-206848P.
 XX PR 26-MAY-2000; 2000US-207727P.
 XX PR 23-OCT-2000; 2000US-242578P.
 XX PR 27-NOV-2000; 2000US-253625P.
 XX PR 22-DEC-2000; 2000US-257931P.
 XX PR 16-FEB-2001; 2001US-269308P.
 XX PA (ELIT-) ELITRA PHARM INC.
 XX PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX DR WPI; 2001-611495/70.
 XX DR N-PSDB; AAS54678.
 XX PT New polynucleotides for the identification and development of
 XX PT antibiotics, comprise sequences of antisense nucleic acids -
 XX PS Example 3; Seq ID No 12412; 51lpp; English.
 XX CC The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 167 AA;

Query Match 41.3%; Score 43; DB 22; Length 167;
 Best Local Similarity 47.4%; Pred. No. 21;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 3 SLSTFFFLFNRSFTQALGK 21
 |::||| ||||| |:
 Db 25 SVATFTLVNRTFTNAQGE 43

RESULT 15
 AAU37429
 ID AAU37429 standard; Protein; 167 AA.
 XX
 AC AAU37429;

XX 14-FEB-2002 (first entry)
 XX Staphylococcus aureus cellular proliferation protein #1599.
 XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.
 XX Staphylococcus aureus.
 OS WO200170955-A2.
 PN 27-SEP-2001.
 XX 21-MAR-2001; 2001WO-US09180.
 XX 21-MAR-2000; 2000US-191078P.
 PR 23-MAY-2000; 2000US-206848P.
 PR 26-MAY-2000; 2000US-207727P.
 PR 23-OCT-2000; 2000US-242578P.
 PR 27-NOV-2000; 2000US-253625P.
 PR 22-DEC-2000; 2000US-257931P.
 PR 16-FEB-2001; 2001US-269308P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
 PI Yamamoto RT, Xu HH;
 XX WPI; 2001-611495/70.
 DR N-PSDB; AAS5288.
 XX New polynucleotides for the identification and development of
 PT antibiotics, comprise sequences of antisense nucleic acids -
 XX Example 3; Seq ID No 13022; 511pp; English.
 XX The invention relates to antisense inhibitors of genes essential to
 CC prokaryotic cellular proliferation, their use in identifying the
 CC genes, their use in the discovery of novel antibiotics, the essential
 CC genes themselves and the encoded proteins. The prokaryotes used are
 CC Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
 CC pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
 CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 167 AA;
 Query Match 41.3%; Score 43; DB 22; Length 167;
 Best Local Similarity 47.4%; Pred. NO. 21;
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
 Qy 3 SLSTFFRLFNRSFTQALGK 21
 Db 25 SVATFTLAVNRTFTNAOGE 43
 Search completed: November 5, 2002, 10:56:08
 Job time : 18.9254 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:03 ; Search time 6.58209 Seconds
(without alignments)
77,929 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 104

Sequence: 1 SGLSTFFRLFNRSFTQALGK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	41.3	479	1	US-08-416-788-2
2	41	39.4	437	1	US-08-764-343-1
3	41	39.4	437	2	US-08-989-925-3
4	40	38.5	82	4	US-08-482-918-59
5	40	38.5	82	4	US-09-224-681-59
6	40	38.5	82	4	US-08-336-728A-59
7	40	38.5	135	4	US-08-936-165A-348
8	40	38.5	164	1	US-08-318-193-77
9	40	38.5	165	2	US-08-955-848A-1
10	40	38.5	165	5	PCT-US95-03866-2
11	40	38.5	166	2	US-08-628-428-2
12	40	38.5	166	2	US-08-628-428-5
13	40	38.5	166	2	US-08-628-428-8
14	40	38.5	166	2	US-09-106-891-2
15	40	38.5	166	4	US-08-172-507-2
16	40	38.5	166	5	PCT-US95-03866-18
17	40	38.5	166	5	PCT-US95-03866-34
18	40	38.5	166	5	PCT-US95-03866-36
19	40	38.5	167	5	PCT-US95-03866-20
20	40	38.5	169	2	US-08-955-848A-2
21	40	38.5	169	2	US-08-955-848A-3
22	40	38.5	169	2	US-08-955-848A-4
23	40	38.5	169	2	US-08-955-848A-5
24	40	38.5	169	2	US-08-955-848A-6
25	40	38.5	169	2	US-08-955-848A-7
26	40	38.5	169	2	US-08-955-848A-8
27	40	38.5	169	2	US-08-955-848A-9

28 40 38.5 169 2 US-08-955-848A-10 Sequence 10, Appl
29 40 38.5 169 2 US-08-955-848A-11 Sequence 11, Appl
30 40 38.5 169 2 US-08-955-848A-12 Sequence 12, Appl
31 40 38.5 169 2 US-08-955-848A-13 Sequence 13, Appl
32 40 38.5 169 2 US-08-955-848A-14 Sequence 14, Appl
33 40 38.5 169 2 US-08-955-848A-15 Sequence 15, Appl
34 40 38.5 169 2 US-08-955-848A-16 Sequence 16, Appl
35 40 38.5 169 2 US-08-955-848A-17 Sequence 17, Appl
36 40 38.5 169 2 US-08-955-848A-18 Sequence 18, Appl
37 40 38.5 169 2 US-08-955-848A-19 Sequence 19, Appl
38 40 38.5 169 2 US-08-955-848A-20 Sequence 20, Appl
39 40 38.5 169 2 US-08-955-848A-21 Sequence 21, Appl
40 40 38.5 169 2 US-08-955-848A-22 Sequence 22, Appl
41 40 38.5 169 2 US-08-955-848A-23 Sequence 23, Appl
42 40 38.5 169 2 US-08-955-848A-24 Sequence 24, Appl
43 40 38.5 169 2 US-08-955-848A-25 Sequence 25, Appl
44 40 38.5 169 2 US-08-955-848A-26 Sequence 26, Appl
45 40 38.5 169 2 US-08-955-848A-27 Sequence 27, Appl

ALIGNMENTS

RESULT 1
US-08-416-788-2
; Sequence 2, Application US/08416788
; Patent No. 5780245
; GENERAL INFORMATION:
; APPLICANT: Maroteaux, Luc
; TITLE OF INVENTION: No. 5780245el Polypeptides Having a Serotonin
; TITLE OF INVENTION: Receptor Activity, Nucleic Acids Coding for These
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426-0107

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/416,788
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/01012
FILING DATE: 13-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92-12280
FILING DATE: 14-OCT-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K.
REGISTRATION NUMBER: 38, 619
REFERENCE/DOCKET NUMBER: EX92008-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 479 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-416-788-2

Query Match 41.3%; Score 43; DB 1; Length 479;
Best Local Similarity 38.1%; Pred. No. 12;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 GSLSTFFRLNRSFTQALGK 21
 || : ||| : || :
 Db 372 SGVNPLIYTLFNKTFREAFGR 392

RESULT 2
 US-08-764-343-1
 ; Sequence 1, Application US/08764343
 ; Patent No. 5739009
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Hawkins, Phillip R.
 ; TITLE OF INVENTION: NOVEL ADIPOCYTE-SPECIFIC
 ; TITLE OF INVENTION: DIFFERENTIATION-RELATED PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/764,343
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0167 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: Consensus

US-08-764-343-1
 Query Match 39.4%; Score 41; DB 1; Length 437;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 10; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

Qy 2 GSLSTFFRLNRSFTQALGK 21
 ||||| : || : || :
 Db 221 GSLST--KLHGRAYOQALSR 238

RESULT 3
 US-08-989-925-3
 ; Sequence 3, Application US/08989925
 ; Patent No. 5989820
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Shah, Purvi
 ; APPLICANT: Corley, Neil C.
 ; TITLE OF INVENTION: HUMAN ADIPOPHILIN-LIKE PROTEIN
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:

Qy 2 GSLSTFFRLNRSFTQALGK 21
 ||||| : || : || :
 Db 221 GSLST--KLHGRAYOQALSR 238

US-08-989-925-3
 Query Match 39.4%; Score 41; DB 1; Length 437;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 10; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

Qy 2 GSLSTFFRLNRSFTQALGK 21
 ||||| : || : || :
 Db 221 GSLST--KLHGRAYOQALSR 238

ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/989,925
 ; FILING DATE: Herewith

CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0440 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166

INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 437 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 1806040

US-08-989-925-3
 Query Match 39.4%; Score 41; DB 2; Length 437;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 10; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

Qy 2 GSLSTFFRLNRSFTQALGK 21
 ||||| : || : || :
 Db 221 GSLST--KLHGRAYOQALSR 238

RESULT 4
 US-08-482-918-59
 ; Sequence 59, Application US/08482918
 ; Patent No. 6207417
 ; GENERAL INFORMATION:
 ; APPLICANT: Zsebo, Krisztina M.
 ; APPLICANT: Bosselman, Robert A.
 ; APPLICANT: Suggs, Sidney V.
 ; APPLICANT: Martin, Francis H.
 ; TITLE OF INVENTION: Stem Cell Factor
 ; NUMBER OF SEQUENCES: 104
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/482,918
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 424

US-08-482-918-59
 Query Match 39.4%; Score 41; DB 2; Length 437;
 Best Local Similarity 50.0%; Pred. No. 24;
 Matches 10; Conservative 5; Mismatches 3; Indels 2; Gaps 1;

Qy 2 GSLSTFFRLNRSFTQALGK 21
 ||||| : || : || :
 Db 221 GSLST--KLHGRAYOQALSR 238

ATTORNEY/AGENT INFORMATION:
 NAME: Clough, David W.
 REGISTRATION NUMBER: 36,107
 REFERENCE/DOCKET NUMBER: 01017/33005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX: 25-3856
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 82 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-482-918-59

Query Match 38.5%; Score 40; DB 4; Length 82;
 Best Local Similarity 87.5%; Pred. No. 5.1;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEPNS 14
 DB 2 FFRIFNRS 9

RESULT 5
 US-09-224-681-59
 Sequence 59, Application US/09224681
 Patent No. 6207454
 GENERAL INFORMATION:
 APPLICANT: Zsebo, Krisztina M.
 APPLICANT: Bosseman, Robert A.
 APPLICANT: Suggs, Sidney V.
 APPLICANT: Martin, Francis H.
 TITLE OF INVENTION: Method for Enhancing the Efficiency of Gene
 Transfer with Stem Cell Factor (SCF) Polypeptide
 NUMBER OF SEQUENCES: 104
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/224,681
 FILING DATE:

CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 09/005,893
 FILING DATE: 12-JAN-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/449,653
 FILING DATE: 24-MAY-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/982,255
 FILING DATE: 25-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/589,701
 FILING DATE: 01-OCT-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/573,616
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/537,198
 FILING DATE: 11-JUN-1990

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/422,383
 FILING DATE: 16-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Clough, David W.
 REGISTRATION NUMBER: 36,107
 REFERENCE/DOCKET NUMBER: 01017/35199
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/474-6300
 TELEFAX: 312/474-0448
 TELEX:
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 82 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-224-681-59

Query Match 38.5%; Score 40; DB 4; Length 82;
 Best Local Similarity 87.5%; Pred. No. 5.1;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEPNS 14
 DB 2 FFRIFNRS 9

RESULT 6
 US-08-336-728A-59
 Sequence 59, Application US/08336728A
 Patent No. 6207802
 GENERAL INFORMATION:
 APPLICANT: Zsebo, Krisztina M.
 APPLICANT: Bosseman, Robert A.
 APPLICANT: Suggs, Sidney V.
 APPLICANT: Martin, Francis H.
 TITLE OF INVENTION: Stem Cell Factor
 NUMBER OF SEQUENCES: 104
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 STREET: 6300 Sears Tower, 233 South Wacker Drive
 CITY: Chicago
 STATE: Illinois
 COUNTRY: United States of America
 ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/336,728A
 FILING DATE: 09-NOV-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/982,255
 FILING DATE: 25-NOV-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/589,701
 FILING DATE: 01-OCT-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/573,616
 FILING DATE: 24-AUG-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/537,198
 FILING DATE: 11-JUN-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/422,383
 FILING DATE: 16-OCT-1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Clough, David W.
 REGISTRATION NUMBER: 36,107

REFERENCE/DOCKET NUMBER: 01017/32956
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-728A-59

Query Match 38.5%; Score 40; DB 4; Length 82;
Best Local Similarity 87.5%; Pred. No. 5.1;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLFNRS 14
| | | | |
DB 2 FFRLFNRS 9

RESULT 7

US-08-936-165A-348
; Sequence 348, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John
; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582el Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50549
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 348:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 135 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

MOLECULE TYPE: Protein
US-08-936-165A-348

Query Match 38.5%; Score 40; DB 4; Length 135;
Best Local Similarity 47.4%; Pred. No. 9.1;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 SLSTFFRLFNRSFTOALGK 21
| | | | | | | | | |
DB 19 SXTTFTIARNRTFTNAQGE 37

RESULT 8

US-08-318-193-77
; Sequence 77, Application US/08318193
; Patent No. 5641663
; GENERAL INFORMATION:
; APPLICANT: GARVIN, Robert T.
; APPLICANT: MALEK, Lawrence T.
; TITLE OF INVENTION: AN EXPRESSION SYSTEM FOR THE SECRETION
; TITLE OF INVENTION: OF BIOACTIVE HUMAN GRANULOCYTE MACROPHAGE COLONY
; TITLE OF INVENTION: STIMULATING FACTOR (GM-CSF) AND OTHER HETEROLOGOUS
; TITLE OF INVENTION: PROTEINS FROM STREPTOMYCES
; NUMBER OF SEQUENCES: 91
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,193
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,314
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 18740/116 CACO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 77:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-318-193-77

Query Match 38.5%; Score 40; DB 1; Length 164;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLFNRS 14
| | | | |
DB 115 FFRLFNRS 122

RESULT 9
US-08-955-848A-1
; Sequence 1, Application US/08955848A
; Patent No. 5969105

GENERAL INFORMATION:
APPLICANT: Mc Wherter, Charles
APPLICANT: Feng, Yiqing
TITLE OF INVENTION: No. 5960105el Stem Cell Factor Receptor
TITLE OF INVENTION: Agonists
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: G. D. Searle & Co.
STREET: P.O. Box 5110
CITY: Chicago
STATE: IL
COUNTRY: U. S. A.
ZIP: 60680
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,848A
FILING DATE: 21-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/029,165
FILING DATE: 25-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Dennis A.
REGISTRATION NUMBER: 34,547
REFERENCE/DOCKET NUMBER: C-2992/1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-737-6986
TELEFAX: 314-737-6972
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-955-848A-1

Query Match 38.5%; Score 40; DB 2; Length 165;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENRS 14
Db 115 FFRIFNRS 122

RESULT 10
PCT-US95-03866-2
Sequence 2, Application PC/TUS9503866
GENERAL INFORMATION:
APPLICANT: Cytomed, Inc. (all states except US)
APPLICANT: Nocke, Karl (US only)
APPLICANT: Lobell, Robert B (US only)
TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
TITLE OF INVENTION: FIT-3/FLK-2 LIGAND
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Neave
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/03866
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/220,379
FILING DATE: 28-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr, James F
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: Cytomed/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 165 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-03866-2

Query Match 38.5%; Score 40; DB 5; Length 165;
Best Local Similarity 87.5%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENRS 14
Db 115 FFRIFNRS 122

RESULT 11
US-08-628-428-2
Sequence 2, Application US/08628428
Patent No. 5885962
GENERAL INFORMATION:
APPLICANT: Lu, Hsieng
TITLE OF INVENTION: SCF ANALOG COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 DeHavilland Drive
CITY: Thousand Oaks
STATE: CA
COUNTRY: USA
ZIP: 91320-1789
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,428
FILING DATE: 05-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Knight, Matthew W
REGISTRATION NUMBER: 36,846
REFERENCE/DOCKET NUMBER: A-400
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..166
OTHER INFORMATION: /note= "Met sequence starts at -1
OTHER INFORMATION: on Sequence No. 5885962 2."
Patent No. 5885962
US-08-628-428-2

Query Match 38.5%; Score 40; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 14
DB 116 FFRIFNR 123

RESULT 12

US-08-628-428-5
; Sequence 5, Application US/08628428
; Patent No. 5885962
; GENERAL INFORMATION:
; APPLICANT: Lu, Hsieng
; TITLE OF INVENTION: SCF ANALOG COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: CA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 05-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: A-400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..166
; OTHER INFORMATION: /note= "Met sequence starts at -1
; OTHER INFORMATION: on Sequence No. 5885962 5."
; Patent No. 5885962
US-08-628-428-5

Query Match 38.5%; Score 40; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 14
DB 116 FFRIFNR 123

RESULT 13

US-08-628-428-8
; Sequence 8, Application US/08628428
; Patent No. 5885962
; GENERAL INFORMATION:
; APPLICANT: Lu, Hsieng
; TITLE OF INVENTION: SCF ANALOG COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks

; STATE: CA
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/628,428
; FILING DATE: 05-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Knight, Matthew W
; REGISTRATION NUMBER: 36,846
; REFERENCE/DOCKET NUMBER: A-400
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: protein
; LOCATION: 1..166
; OTHER INFORMATION: /note= "Met sequence starts at -1
; OTHER INFORMATION: on Sequence No. 5885962 8."
; Patent No. 5885962
US-08-628-428-8

Query Match 38.5%; Score 40; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 14
DB 116 FFRIFNR 123

RESULT 14

US-09-106-891-2
; Sequence 2, Application US/09106891
; Patent No. 5965522
; GENERAL INFORMATION:
; APPLICANT: Herhenson, Susan
; TITLE OF INVENTION: No. 5965522el Stem Cell Factor Formulations and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 DeHavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/106,891
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/172,507
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pessin, Karol M.
; REFERENCE/DOCKET NUMBER: A-276
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

LENGTH: 166 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-106-891-2

Query Match 38.5%; Score 40; DB 2; Length 166;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEFRS 14
Db 116 FFRIFNRS 123

RESULT 15

US-08-172-507-2
Sequence 2, Application US/08172507
Patent No. 6288030
GENERAL INFORMATION:
APPLICANT: Hershenson, Susan
TITLE OF INVENTION: No. 6288030el Stem Cell Factor Formulations and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Inc.
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/172,507
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Pessin, Karol M.
REFERENCE/DOCKET NUMBER: A-276
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 166 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-172-507-2

Query Match 38.5%; Score 40; DB 4; Length 166;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEFRS 14
Db 116 FFRIFNRS 123

Search completed: November 5, 2002, 10:58:19
Job time : 6.58209 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run On: November 5, 2002, 10:55:02 ; Search time 8.14925 Seconds
(without alignments)
247.615 Million cell updates/sec

Title: US-09-833-017b-4

Perfect score: 104

Sequence: 1 SGLSTFFRLNRSFTQALGK 21

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_71.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	48	46.2	1083	2 S76111	acriflavin resista
2	45	43.3	124	2 T51577	gene wnt-7C protel
3	45	43.3	850	2 P95260	ABC transporter, p
4	45	43.3	850	2 A98126	conserved hypothet
5	44.5	42.8	689	1 S29621	UDPGlucose 4-epime
6	44	42.3	490	2 I48163	cytochrome P450 -
7	44	42.3	889	2 H84506	probable retroelem
8	44	42.3	970	2 G84939	hypothetical prote
9	43	41.3	167	2 G89802	hypothetical prote
10	43	41.3	230	2 S46737	hypothetical prote
11	43	41.3	305	2 P69748	hypothetical prote
12	43	41.3	412	2 C38351	phosphoprotein pho
13	43	41.3	443	2 B38351	phosphoprotein pho
14	43	41.3	443	2 S65685	protein phosphatas
15	43	41.3	447	2 A38351	phosphoprotein pho
16	43	41.3	447	2 A41805	phosphoprotein pho
17	43	41.3	468	2 JC5417	phosphoprotein pho
18	43	41.3	468	2 S65951	[phosphorylase] ph
19	43	41.3	479	2 S23562	serotonin receptor
20	43	41.3	504	2 S27269	minor inner core p
21	43	41.3	835	1 P3XR90	hypothetical prote
22	43	41.3	1123	2 D36790	opuaa (AF234619) [
23	42.5	40.9	361	2 D98243	hypothetical prote
24	42.5	40.9	361	2 AH3042	hypothetical prote
25	42	40.4	113	2 T09627	positive transcrip
26	42	40.4	139	2 A71190	hypothetical prote
27	42	40.4	155	2 T31981	hypothetical prote
28	42	40.4	206	2 A81231	50S ribosomal prot
29	42	40.4	393	2 B82283	conserved hypothet

30 42 40.4 490 2 I48162 cytochrome P450 -

31 42 40.4 490 2 I48189 cytochrome P450 II

32 42 40.4 499 2 A45778 phosphoprotein pho

33 42 40.4 709 2 S73497 cell division prot

34 42 40.4 851 2 A83484 probable heme util

35 41.5 39.9 1242 2 S51246 probable DNA repai

36 41 39.4 136 2 T00295 probable transposa

37 41 39.4 158 2 D71111 hypothetical prote

38 41 39.4 253 2 S70367 stem cell factor s

39 41 39.4 262 2 P83546 probable transcrip

40 41 39.4 265 1 EVECIT1 istB protein - Esc

41 41 39.4 287 2 S70366 stem cell factor 1

42 41 39.4 373 2 T47150 hypothetical prote

43 41 39.4 435 2 E82848 alpha-ketoglutarat

44 41 39.4 527 2 T39927 probable serine-ri

45 41 39.4 556 2 T39476 probable serine-ri

ALIGNMENTS

RESULT 1

S76111

acriflavin resistance protein envD - Synechocystis sp. (strain PCC 6803)

N:Alternate names: protein slr0369

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

R:Accession: S76111

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S76111

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1083 <KAN>

A:Cross-references: EMBL:D63999; GB:AB001339; NID:gl001396; PIDN:BAAL0089.1; PID:gl00

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: envD

C:Superfamily: acriflavin resistance protein

Query Match 46.2%; Score 48; DB 2; Length 1083;

Best Local Similarity 64.3%; Pred. No. 10; 3; Indels 0; Gaps 0;

Matches 9; Conservative 2; Mismatches

QY 2 GSLSTFFRLNRSF 15

DB 514 GPLAWFFNLFNRKF 527

RESULT 2

I51577

gene wnt-7C protein - African clawed frog (fragment)

C:Species: Xenopus laevis (African clawed frog)

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C:Accession: I51577

R:Wolda, S.L.; Moon, R.T.

Oncogene 7, 1941-1947, 1992

A:Title: Cloning and developmental expression in Xenopus laevis of seven additional m

A:Reference number: I51571; MUID:93026368

A:Accession: I51577

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-124 <WOI>

A:Cross-references: GB:L07535; NID:g214889; PIDN:AAA49988.1; PID:g214890

C:Genetics:

A:Gene: wnt-7C

C:Superfamily: int-1 transforming protein

```
Query Match 43.3%; Score 45; DB 2; Length 124;
Best Local Similarity 50.0%; Pred. No. 3;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 SGSLSSTFFRLNRSFTQA 18
   :||: | : ||: |||
DB 81 TGSIGTYGRFCNRTSTQA 98

RESULT 3
F95260
A: C: Species: Streptococcus pneumoniae
A: C: Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
A: C: Accession: F95260
A: C: Superfamily: H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heid
on, J.D.; Unayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
son, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A: Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison,
A: Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A: Reference number: A95000; MUID:21357209; PMID:11463916
A: Accession: F95260
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-850 <KUR>
A: Cross-references: GB:AE005672; PIDN:AAK76279.1; PID:g14973742; GSPDB:GN00164; TIGR:SP4
A: Experimental source: strain TIGR4
A: Genetics:
A: Gene: SP2231

Query Match 43.3%; Score 45; DB 2; Length 850;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 TFFRLNRSFTQALCK 21
   ||:| | :||:| :|
DB 733 TFYRLDTKTFTFAIQK 748

RESULT 4
A98126
A: C: Species: conserved hypothetical protein spr2036 [imported] - Streptococcus pneumoniae (strain R6)
A: C: Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
A: C: Accession: A98126
A: C: Superfamily: Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; B
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A: Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A: Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A: Reference number: A97872; MUID:21429245; PMID:11544234
A: Accession: A98126
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-850 <KUR>
A: Cross-references: GB:AE007317; PIDN:AA00838.1; PID:g15459743; GSPDB:GN00174
A: Genetics:
A: Gene: spr2036

Query Match 43.3%; Score 45; DB 2; Length 850;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 TFFRLNRSFTQALCK 21
   ||:| | :||:| :|
DB 733 TFYRLDTKTFTFAIQK 748

RESULT 5
S29621
A: C: Species: UDPglucose 4-epimerase (BC 5.1.3.2) - yeast (Pachysolen tannophilus)
A: C: Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
A: C: Accession: S29621
A: C: Superfamily: R; Skrzypek, M.; Maleszka, R.
submitted to the EMBL data Library, October 1992
A: Description: Cloning and sequencing of the UDP-galactose-4-epimerase gene from Pach
A: Reference number: S29621
A: Accession: S29621
A: Molecule type: DNA
A: Residues: 1-689 <SKR>
A: Cross-references: EMBL:X68593; NID:g3264; PIDN:CAA48580.1; PID:g3265
A: Genetics:
A: Gene: GALI0
A: Superfamily: yeast UDPglucose 4-epimerase; UDPglucose 4-epimerase homology
C: Keywords: galactose metabolism; isomerase
F: 4-343/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 42.8%; Score 44.5; DB 1; Length 689;
Best Local Similarity 47.6%; Pred. No. 24;
Matches 10; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

QY 1 SGSLSSTFFRLNRSFTQALCK 21
   :||: | : ||: |||
DB 270 TGGSTVFEVEN-AFCRAVGK 289

RESULT 6
I48163
A: C: Species: cytochrome P450 - golden hamster
A: C: Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 04-Mar-2000
A: C: Accession: I48163
A: C: Superfamily: R; Sakuma, T.; Masaki, K.; Itoh, S.; Yokoi, T.; Kamataki, T.
Mol. Pharmacol. 45, 228-236, 1994
A: Title: Sex-related difference in the expression of cytochrome P450 in hamsters: cDN
A: Reference number: I48162; MUID:94158799
A: Accession: I48163
A: Status: preliminary; translated from GB/EMBL/DBJ
A: Molecule type: mRNA
A: Residues: 1-490 <RES>
A: Cross-references: GB:D11436; NID:g220312; PIDN:BAA02002.1; PID:g220313
A: Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C: Keywords: chromoprotein; heme; iron; metalloprotein; transmembrane protein
F: 294-457/Domain: cytochrome P450 homology <P45>
F: 435/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 42.3%; Score 44; DB 2; Length 490;
Best Local Similarity 42.9%; Pred. No. 21;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 SGSLSSTFFRLNRSFTQALCK 21
   ||||:| :||:| :|
DB 51 SGSLSSTFFRLNRSFTQALCK 71

RESULT 7
H84506
A: C: Species: probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
A: C: Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
A: C: Accession: H84506
A: C: Superfamily: R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter
Nature 402, 761-768, 1999
A: Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A: Reference number: A84420; MUID:20083487
A: Accession: H84506
A: Status: preliminary
A: Molecule type: DNA
A: Residues: 1-889 <STO>
A: Cross-references: GB:AE000209; NID:g4417309; PIDN:AAD0433.1; GSPDB:GN00139
```

RESULT 10

R:Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W.
Biochemistry 30, 3589-3597, 1991
A:Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A:Reference number: A38351; MUID:91198016

A:Accession: C38351
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-412 <MAV>
A:Cross-references: GB:J05328
C:Keywords: phosphoric monoester hydrolase

Query Match 41.3%; Score 43; DB 2; Length 412;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLSTFFFLFNRS 14
:||: |||:|:|:
Db 324 TGSYNFFRFDNRN 337

RESULT 13

B38351
phosphoprotein phosphatase 2-beta regulatory chain - human
N:Alternate names: phosphoprotein phosphatase 2A-beta 55K regulatory chain B
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 05-Nov-1999
C:Accession: B38351

R:Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W.
Biochemistry 30, 3589-3597, 1991
A:Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A:Reference number: A38351; MUID:91198016

A:Accession: B38351
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-443 <MAV>
A:Cross-references: GB:M64930; GB:J05328; NID:gl90423; PIDN:AAA36493.1; PID:gl90426
C:Genetics:
A:Gene: GDB:PPP2R2B
A:Cross-references: GDB:136801

Query Match 41.3%; Score 43; DB 2; Length 443;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLSTFFFLFNRS 14
:||: |||:|:|:
Db 355 TGSYNFFRFDNRN 368

RESULT 14

S65685
protein phosphatase 2A, chain B (BR beta) - rat
C:Species: Rattus sp. (rat)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 05-Nov-1999
C:Accession: S65685
R:Akiyama, N.; Shima, H.; Hatano, Y.; Osawa, Y.; Sugimura, T.; Nagao, M.
Eur. J. Biochem. 230, 766-772, 1995
A:Title: cDNA cloning of BR-gamma, a novel brain-specific isoform of the B regulatory subunit of protein phosphatase 2A: evidence
A:Reference number: S65685; MUID:95331316

A:Accession: S65685
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-443 <AKI>
A:Cross-references: EMBL:D38260; NID:gl065605; PIDN:BA07412.1; PID:dl007991; PID:gl7774

Query Match 41.3%; Score 43; DB 2; Length 443;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLSTFFFLFNRS 14
:||: |||:|:|:
Db 355 TGSYNFFRFDNRN 368

RESULT 15

A38351
phosphoprotein phosphatase 2-alpha regulatory chain - human
N:Alternate names: phosphoprotein phosphatase 2A-alpha 55K regulatory chain B
C:Species: Homo sapiens (man)
C:Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 05-Nov-1999
C:Accession: A38351
R:Mayer, R.E.; Hendrix, P.; Cron, P.; Matthies, R.; Stone, S.R.; Goris, J.; Merlevede, W.
Biochemistry 30, 3589-3597, 1991
A:Title: Structure of the 55-kDa regulatory subunit of protein phosphatase 2A: evidence
A:Reference number: A38351; MUID:91198016

A:Accession: A38351
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-447 <MAV>
A:Cross-references: GB:M64929; GB:J05328; NID:gl90421; PIDN:AAA36490.1; PID:gl90422
C:Genetics:
A:Gene: GDB:PPP2R2A
A:Cross-references: GDB:136800

Query Match 41.3%; Score 43; DB 2; Length 447;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SGLSTFFFLFNRS 14
:||: |||:|:|:
Db 359 TGSYNFFRFDNRN 372

Search completed: November 5, 2002, 10:56:44
Job time: 12.1493 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:03 ; Search time 5.01493 Seconds
(without alignments)
162.138 Million cell updates/sec

Title: US-09-833-017B-4
Perfect score: 104
Sequence: 1 SGSLSTFFFLNRSTQALGK 21

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	45	43.3	135	1 WN7C_XENLA	P31290 xenopus lae
2	44.5	42.8	689	1 GALX_PACTA	P40801 pachysolen
3	44	42.3	490	1 CPCP_MESAU	P33264 mesocricetu
4	44	42.3	934	1 ORP6_HUMAN	Q9bzf3 homo sapien
5	44	42.3	970	1 Y087_BUCAL	P57189 buchiera ap
6	43	41.3	230	1 FILL_YEAST	P38771 saccharomyc
7	43	41.3	413	1 ZABG_RABIT	Q00006 o serine/th
8	43	41.3	426	1 ZABA_PIG	Q29090 s serine/th
9	43	41.3	443	1 ZABB_HUMAN	Q00005 homo sapien
10	43	41.3	443	1 ZABB_PIG	P54614 sus scrofa
11	43	41.3	443	1 ZABB_RAT	P36877 h serine/th
12	43	41.3	447	1 ZABA_HUMAN	Q00007 h serine/th
13	43	41.3	447	1 ZABA_RAT	P36876 r serine/th
14	43	41.3	453	1 ZABD_RAT	P56932 r serine/th
15	43	41.3	479	1 ZABE_RAT	P30994 rattus norv
16	43	41.3	504	1 SH2B_MOUSE	Q02152 mus musculu
17	43	41.3	835	1 VP3_ROTSI	P15736 simian il r
18	43	41.3	1133	1 VG39_HSVII	Q00143 ictaluriid h
19	42	40.4	113	1 RAMA_KLEPN	Q48413 klebsiella
20	42	40.4	490	1 CPCP_MESAU	Q08078 mesocricetu
21	42	40.4	490	1 CPCQ_MESAU	P33263 mesocricetu
22	42	40.4	499	1 ZABA_DROME	P36872 drosophila
23	42	40.4	709	1 FTSH_MYCPN	P75120 mycoplasma
24	41.5	39.9	1242	1 MSU6_YEAST	Q03834 saccharomyc
25	41	39.4	113	1 RAMA_ENTCL	P55922 enterobacte
26	41	39.4	189	1 APOD_MOUSE	P51910 mus musculu
27	41	39.4	265	1 ISTB_PSEAE	P15026 pseudomonas
28	41	39.4	437	1 ADFP_HUMAN	Q99541 homo sapien
29	41	39.4	540	1 GUXI_ASFAC	O59843 aspergillus
30	41	39.4	2436	1 ABC2_HUMAN	Q9bzc7 homo sapien
31	40	38.5	96	1 SH2B_CAVPO	P97267 cavia porce
32	40	38.5	267	1 Y084_DEIRA	Q9irk5 deinococcus
33	40	38.5	273	1 SCF_HUMAN	P21583 homo sapien

Q06220 canis famil
Q9y2t4 h serine/th
P50410 o serine/th
P97888 r serine/th
P41595 homo sapien
Q09299 caenorhabdi
Q10947 caenorhabdi
P19812 saccharomyc
P27427 dhori virus
Q03518 homo sapien
P24805 nicotiana t
P01365 saccharomyc

ALIGNMENTS

RESULT 1
WN7C_XENLA STANDARD; PRT; 135 AA.
AC P31290;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE WNT-7C protein (XWNT-7C) (Fragment).
GN WNT-7C.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93026368; PubMed=1408135;
RA Wolda S.L., Moon R.T.;
RT "Cloning and developmental expression in Xenopus laevis of seven additional members of the Wnt family."
RL Oncogene 7:1941-1947(1992).
CC -!- FUNCTION: LIGAND FOR MEMBERS OF THE FRIZZLED FAMILY OF SEVEN TRANSMEMBRANE RECEPTORS. PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
CC -!- SUBCELLULAR LOCATION: Possibly secreted and associates with the extracellular matrix.
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; L07535; AAA49988.1; -
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt1; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein.
FT NON_TER 1
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC...) (POTENTIAL).
FT NON_TER 135 135 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 135 AA; 15344 MW; 8456A5DA03A39B4C CRC64;

Query Match 43.3%; Score 45; DB 1; Length 135;
Best Local Similarity 50.0%; Pred. No. 1.5; Indels 0; Gaps 0;
Matches 9; Conservative 4; Mismatches 5;
QY 1 SGSLSTFFFLNRSTQ 18

OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21376257; PubMed=11483621;
 RA Lehto M., Laitinen S., Chinetti G.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuho Y., Sakaki N.;
 RT "NEO human cDNA sequencing project."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE OSBP FAMILY.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AF323728; AAC53409.1; -
 DR EMBL; AK027600; BAB5223.1; -
 DR InterPro: IPR000648; Oxyesterol_BP.
 DR Pfam: PF01237; Oxyesterol_BP; 1.
 DR Pfam: PF00169; PH; 1.
 DR SMART: SM00233; PH; 1.
 DR PROSITE: PS01013; OSBP; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 KW Lipid transport; Transport.
 FT DOMAIN 86 181 PH.
 FT CONFLICT 319 319 S-> G (IN REF. 2).
 FT CONFLICT 714 714 N-> S (IN REF. 2).
 SQ SEQUENCE 934 AA; 106305 MW; D20F90EA34C81497 CRC64;
 Query Match 42.3%; Score 44; DB 1; Length 934;
 Best Local Similarity 42.9%; Pred. No. 17;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 1 SGSLSTFFFLFNRSTQALGK 21
 DB 634 SGYCTYFRAGSKPFPNVIGE 654
 RESULT 5
 Y087_BUCAI STANDARD; PRT; 970 AA.
 AC P57189;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein BU087.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 OS symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TOKYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 RT Buchnera sp. APS."
 RL Nature 407:81-86(2000).
 CC -!- SUBUNIT: SOME, TO E.COLI YTFN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AP001118; BAB12807.1; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 12 32 POTENTIAL.
 SQ SEQUENCE 970 AA; 114477 MW; 16B7BADB129F422E CRC64;
 Query Match 42.3%; Score 44; DB 1; Length 970;
 Best Local Similarity 80.0%; Pred. No. 17;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 6 TFFRLFNRSF 15
 DB 703 SFENLFRSE 712
 RESULT 6
 FILL_YEAST STANDARD; PRT; 230 AA.
 AC P38771; 1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE FILL protein, mitochondrial precursor.
 GN FILL OR KIM4 OR YHR038W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MT8-1;
 RX MEDLINE=98417448; PubMed=9746366;
 RA Kanai T., Takeshita S., Atomi H., Umemura K., Ueda M., Tanaka A.;
 RT "A regulatory factor, FILLP, involved in derepression of the
 RT isocitrate lyase gene in Saccharomyces cerevisiae -- a possible
 RT mitochondrial protein necessary for protein synthesis in
 RT mitochondria".
 RL Eur. J. Biochem. 256:212-220(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RX MEDLINE=94378003; PubMed=8091229;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
 Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
 Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
 Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
 Nham M., Rifkin L., Riles L., St Peter H., Trevisan E., Vaughan K.,
 Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
 Vaudin M.;
 RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
 RT VIII."
 RL Science 265:2077-2082(1994).
 RN [3]
 RP SIMILARITY TO RRF.
 RX MEDLINE=96155086; PubMed=8563640;
 RA Ouzounis C., Bork P., Casari G., Sander C.;
 RT "New protein functions in yeast chromosome VIII."
 RL Protein sci. 4:2424-2428(1995).
 CC -!- FUNCTION: INVOLVED IN GLUCOSE REPRESSION/DEREPRESSION. MAY BE
 CC NECESSARY FOR PROTEIN SYNTHESIS IN MITOCHONDRIA.
 CC -!- SUBCELLULAR LOCATION: Mitochondrial.
 CC -!- SIMILARITY: BELONGS TO THE RRF FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB016033; BAA31687.1; -;
 DR EMBL; U00062; AAB68906.1; -;
 DR PIR; S46737; S46737;
 DR SGD; S0001080; FILL;
 DR InterPro; IPR002661; RRF;
 DR Pfam; PF01765; RRF; 1;
 DR Protein biosynthesis; Transist peptide; Mitochondrion.
 FT TRANSIT 1 24 MITOCHONDRION (POTENTIAL).
 FT CHAIN 25 230 FILL PROTEIN.
 SQ SEQUENCE 230 AA; 26406 MW; 9CBDC8C5F8F3008 CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 230;
 Best Local Similarity 80.0%; Pred. No. 5.6;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 OY 9 RLNRSTQA 18
 DB 17 RLNRSTQA 26
 RESULT 7
 ID 2ABB_RABIT STANDARD; PRT; 413 AA.
 AC Q00006;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B, B55-
 DE beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
 DE R2-beta isoform) (Fragment).
 GN PPP2R2B.
 OS Oryctolagus cuniculus (Rabbit).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 CC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=91198015; PubMed=1849734;
 RA Mayer R.E., Hendrix P., Cron P., Matthies R., Stone S.R.,
 RA Goris J., Merlieve W., Holsinger J., Hemmings B.A.;
 RT "Structure of the 55-kDa regulatory subunit of protein phosphatase
 RT 2A: evidence for a neuronal-specific isoform.";
 RL Biochemistry 30:3589-3597(1991).
 CC -!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M64931; AAA31458.1; -;
 DR InterPro; IPR000009; PP2A_PR55.
 DR Pfam; PF00400; WD40; 5;
 DR SMART; SM00320; WD40; 2;
 DR PROSITE; PS01024; PR55_1; 1;
 DR PROSITE; PS01025; PR55_2; 1;
 KW Multigene family.
 FT NON_TER 1
 SQ SEQUENCE 413 AA; 48243 MW; 49237B7817EB8FE2 CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 413;
 Best Local Similarity 50.0%; Pred. No. 10;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SGLSTFFRLNRS 14
 DB 325 TGSYNFFRFDNR 338
 RESULT 8
 ID 2ABA_PIG STANDARD; PRT; 426 AA.
 AC Q29090;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE alpha isoform (PP2A, subunit B, B-alpha isoform) (PP2A, subunit B,
 DE B55-alpha isoform) (PP2A, subunit B, PR55-alpha isoform) (PP2A,
 DE subunit B, R2-alpha isoform) (Fragment).
 GN PPP2R2A.
 OS Sus scrofa (Pig).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 CC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mayer-Jackel R.E.;
 RL Thesis (1992), Friedrich Miescher Institut / Basel, Switzerland.
 CC -!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z34932; CAAB4404.1; -;
 DR InterPro; IPR000009; PP2A_PR55.
 DR Pfam; PF00400; WD40; 5;
 DR PRINTS; SM00600; PP2APR55.
 DR SMART; SM00320; WD40; 3;
 DR PROSITE; PS01024; PR55_1; 1;
 DR PROSITE; PS01025; PR55_2; 1;

KW Multigene family.
 SQ NON_TER 1
 SEQUENCE 426 AA; 49613 MW; 3AAD7EB338803334 CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 426;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLSTFFFLFNRS 14
 :||:||||:|:
 Db 338 TGSYNFFRMDN 351

RESULT 9
 2ABB_HUMAN
 ID 2ABB_HUMAN STANDARD; PRT; 443 AA.
 AC Q00005;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B, B55-
 DE beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
 DE R2-beta isoform).
 GN PPP2R2B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tissue=Fetal brain;
 RC MEDLINE=91198016; PubMed=1849734;
 RA Mayer R.E., Hendrix P., Cron P., Matthies R., Stone S.R.,
 RA Goris J., Merlevede W., Hofsteenge J., Hemmings B.A.;
 RT 2A: evidence for a neuronal-specific isoform.";
 RT Biochemistry 30:3589-3597(1991).
 CC !- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC !- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A). THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B''/PR72/PR130/PR59 AND R3/B''/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC !- TISSUE SPECIFICITY: BRAIN.
 CC !- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M64930; AAA36493.1;
 CC PIR: B38351; B38351.
 CC MIM: 604325;
 CC InterPro: IPR000009; PP2A_PR55.
 CC Pfam: PF00400; WD40; 6.
 CC PRINTS: PR00600; PP2APR55.
 CC SMART: SM00320; WD40; 3.
 CC PROSITE: PS01024; PR55_1; 1.
 CC PROSITE: PS01025; PR55_2; 1.
 CC Multigene family.
 KW SEQUENCE 443 AA; 51710 MW; C383C834B2852B8F CRC64;

Query Match 41.3%; Score 43; DB 1; Length 443;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLSTFFFLFNRS 14
 :||:||||:|:
 Db 355 TGSYNFFRMDN 368

RESULT 10
 2ABB_PIG
 ID 2ABB_PIG STANDARD; PRT; 443 AA.
 AC P54614;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE beta isoform (PP2A, subunit B, B-beta isoform) (PP2A, subunit B, B55-
 DE beta isoform) (PP2A, subunit B, PR55-beta isoform) (PP2A, subunit B,
 DE R2-beta isoform).
 GN PPP2R2B.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mayer-Jaekel R.E.;
 RC Thesis (1992); Friedrich Miescher Institut / Basel, Switzerland.
 CC !- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC !- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A). THAT ASSOCIATES
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B''/PR72/PR130/PR59 AND R3/B''/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC !- TISSUE SPECIFICITY: BRAIN.
 CC !- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z34933; CAA84405.1;
 CC InterPro: IPR000009; PP2A_PR55.
 CC Pfam: PF00400; WD40; 5.
 CC PRINTS: PR00600; PP2APR55.
 CC SMART: SM00320; WD40; 4.
 CC PROSITE: PS01024; PR55_1; 1.
 CC PROSITE: PS01025; PR55_2; 1.
 CC Multigene family.
 KW SEQUENCE 443 AA; 51459 MW; P8562FC696719F41 CRC64;

Query Match 41.3%; Score 43; DB 1; Length 443;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGLSTFFFLFNRS 14
 :||:||||:|:
 Db 355 TGSYNFFRMDN 368

Query Match 41.3%; Score 43; DB 1; Length 447;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLFNRS 14
 :||:||||:|:
 Db 359 TGSYNFFRMEDRN 372

RESULT 13
 2ABA_RAT
 ID 2ABA_RAT STANDARD; PRT: 447 AA.
 AC P36876; P36878; O35512;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE alpha isoform (PP2A, subunit B, B-alpha isoform) (PP2A, subunit B,
 DE B55-alpha isoform) (PP2A, subunit B, PR55-alpha isoform) (PP2A,
 DE subunit B, R2-alpha isoform) (PP2A, subunit B, BRA isoform).
 GN PPP2R2A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92114192; PubMed=1370560;
 RA Pallas D.C., Weller W., Jaspers S., Miller T.B. Jr., Lane W.S.,
 RA Roberts T.M.;
 RT "The third subunit of protein phosphatase 2A (PP2A), a 55-kilodalton
 RT protein which is apparently substituted for by T antigens in
 RT complexes with the 36- and 63-kilodalton PP2A subunits, bears little
 RT resemblance to T antigens.";
 RL J. Virol. 66:886-893(1992).
 RN [2]
 RP SEQUENCE OF 80-272 FROM N.A.
 RC STRAIN=FISCHER 344;
 RX MEDLINE=93279382; PubMed=839301;
 RA Hatano Y., Shima H., Haneji T., Miura A.B., Sugimura T., Nagao M.;
 RA "Expression of PP2A B regulatory subunit beta isotype in rat
 RA testis.";
 RL FEBS Lett. 324:71-75(1993).
 CC -!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATE
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M83298; AAA41910.1; -
 CC EMBL; M83297; AAA41909.1; -
 CC EMBL; D14419; AAA21904.1; -
 CC PIR; A41805; A41805.
 CC InterPro; IPR000009; PP2A_PR55.
 CC InterPro; IPR001680; WD40.

DR Pfam; PF00400; WD40; 6.
 DR PRINTS; PRO0600; PP2APR55.
 DR SMART; SM00320; WD40; 3.
 DR PROSITE; PS01024; PR55_1; 1.
 DR DR PROSITE; PS01025; PR55_2; 1.
 KW Multigene family.
 FT VARIANT 60 60 E -> ESPKVAALREASNLMSQ.
 FT CONFLICT 105 105 K -> E (IN REF. 1: AAA41909).
 FT CONFLICT 105 105 K -> R (IN REF. 2).
 FT CONFLICT 213 213 N -> S (IN REF. 2).
 FT CONFLICT 222 222 M -> V (IN REF. 2).
 SQ SEQUENCE 447 AA; 51678 MW; 180ACB37D9DA4ECE CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 447;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLFNRS 14
 :||:||||:|:
 Db 359 TGSYNFFRMEDRN 372

RESULT 14
 2ABD_RAT
 ID 2ABD_RAT STANDARD; PRT: 453 AA.
 AC P56932;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE Serine/threonine protein phosphatase 2A, 55 kDa regulatory subunit B,
 DE delta isoform (PP2A, subunit B, B-delta isoform) (PP2A, subunit B,
 DE B55-delta isoform) (PP2A, subunit B, PR55-delta isoform) (PP2A,
 DE subunit B, R2-delta isoform).
 GN PPP2R2D.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;
 RX MEDLINE=20026081; PubMed=10556517;
 RA Strack S., Chang D., Zaucha J.A., Colbran R.J., Wadzinski B.E.;
 RA "Cloning and characterization of B delta, a novel regulatory subunit
 RA of protein phosphatase 2A.";
 RL FEBS Lett. 460:462-466(1999).
 CC -!- FUNCTION: THE B REGULATORY SUBUNIT MIGHT MODULATE SUBSTRATE
 CC SELECTIVITY AND CATALYTIC ACTIVITY, AND ALSO MIGHT DIRECT THE
 CC LOCALIZATION OF THE CATALYTIC ENZYME TO A PARTICULAR SUBCELLULAR
 CC COMPARTMENT.
 CC -!- SUBUNIT: PP2A CONSISTS OF A COMMON HETERODIMERIC CORE ENZYME,
 CC COMPOSED OF A 36 kDa CATALYTIC SUBUNIT (SUBUNIT C) AND A 65 kDa
 CC CONSTANT REGULATORY SUBUNIT (PR65 OR SUBUNIT A), THAT ASSOCIATE
 CC WITH A VARIETY OF REGULATORY SUBUNITS. PROTEINS THAT ASSOCIATE
 CC WITH THE CORE DIMER INCLUDE THREE FAMILIES OF REGULATORY SUBUNITS
 CC B (THE R2/B/PR55/B55, R3/B'/PR72/PR130/PR59 AND R5/B'/B56
 CC FAMILIES), THE 48 kDa VARIABLE REGULATORY SUBUNIT, VIRAL PROTEINS,
 CC AND CELL SIGNALING MOLECULES.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH HIGH LEVELS IN BRAIN,
 CC HEART, PLACENTA, SKELETAL MUSCLE, TESTIS, THYMUS AND SPLEEN.
 CC -!- SIMILARITY: BELONGS TO THE PHOSPHATASE 2A REGULATORY SUBUNIT B
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF180350; AAF08536.1; -

DR InterPro: IPR000009; PP2A_PR55.
 DR InterPro: IPR001680; WD40.
 DR Pfam: PF00400; WD40; 6.
 DR PRINTS; PR00600; PP2APR55.
 DR SMART; SM00320; WD40; 2.
 KW Multigene family.
 FT DOMAIN 3 8 POLY-GLY
 SQ SEQUENCE 453 AA; 51982 MW; 733E80A93A5BC2BB CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 453;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 OY 1 SGLSLSTFFRLFNRS 14
 :|| : |||:|:|:
 DB 365 TGSYNNEFRMEDRN 378
 RESULT 15
 ID SH2B_RAT STANDARD; PRT; 479 AA.
 AC P30994;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE 5-hydroxytryptamine 2B receptor (5-HT-2B) (Serotonin receptor)
 DE (5-HT-2F) (Stomach fundus serotonin receptor).
 GN HPR2B OR SRL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY;
 RX MEDLINE=92371456; PubMed=1505525;
 RA Roquet M., Hoyer D., Pardo L.A., Parekh A., Kluxen F.-W.,
 RA Kaikman M.O., Stuehmer W., Luebbert H.;
 RT "Cloning and functional characterization of the rat stomach fundus
 serotonin receptor.";
 RL EMBO J. 11:3481-3487(1992).
 CC -1- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR
 CC 5-HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
 CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THIS RECEPTOR
 CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: STOMACH FUNDUS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST TO THE OTHER 5HT-2 SUBTYPE RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X66842; CAA47318.1; -
 CC PIR; S23562; S23562.
 CC GCRDB; GCR_0434; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECIP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECIP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate.
 FT DOMAIN 1 55 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 56 78 1 (POTENTIAL).
 FT DOMAIN 79 90 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 91 112 2 (POTENTIAL).

FT DOMAIN 113 128 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 129 150 3 (POTENTIAL).
 FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 171 191 4 (POTENTIAL).
 FT DOMAIN 192 215 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 216 238 5 (POTENTIAL).
 FT DOMAIN 239 323 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 324 344 6 (POTENTIAL).
 FT DOMAIN 345 359 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 360 382 7 (POTENTIAL).
 FT DOMAIN 383 479 CYTOPLASMIC (POTENTIAL).
 DISULFID 127 206 BY SIMILARITY.
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT LIPID 396 396 PALMITATE (POTENTIAL).
 SQ SEQUENCE 479 AA; 53651 MW; 17FFC73213B42038 CRC64;
 Query Match 41.3%; Score 43; DB 1; Length 479;
 Best Local Similarity 38.1%; Pred. No. 12;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 OY 1 SGLSLSTFFRLFNRSFTQALGK 21
 :|| : |||:|:|:
 DB 372 SGVNPPLYTLLENKTFREAFGR 392

Search completed: November 5, 2002, 10:57:55
 Job time : 8.01493 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:55:02 ; Search time 13.4776 Seconds
(without alignments)
269,550 Million cell updates/sec

Title: US-09-833-017b-4
Perfect score: 104
Sequence: 1 SGSLSTFFRLNRSFTQALGK 21

Scoring table: BLOSUM62
Gapop 10.0, Capext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	46	2 Q99015	Q99015 streptococ
2	104	100.0	46	2 Q9AK7	Q9AK7 streptococ
3	86	82.7	43	2 Q9AK6	Q9AK6 streptococ
4	48	46.2	1083	16 Q5584	Q5584 synochocyst
5	45	43.3	390	11 Q9QZS6	Q9QZS6 mus musculu
6	45	43.3	835	12 Q9QNB1	Q9QNB1 human rotav
7	45	43.3	850	16 Q97N40	Q97N40 streptococ
8	45	43.3	1528	12 Q991N1	Q991N1 little cher
9	45	43.3	1980	12 Q991N2	Q991N2 little cher
10	44	42.3	311	4 Q96RX7	Q96RX7 homo sapien
11	44	42.3	381	4 Q96015	Q96015 homo sapien
12	44	42.3	431	5 Q9N3H3	Q9N3H3 caenorhabdi
13	44	42.3	719	4 Q96SR1	Q96SR1 homo sapien
14	44	42.3	740	6 Q9N007	Q9N007 macaca fasc
15	44	42.3	799	11 Q9DBI0	Q9DBI0 mus musculu
16	44	42.3	844	6 Q9N006	Q9N006 macaca fasc

17	44	42.3	889	10 Q9SK7	Q9SK7 arabidopsis
18	44	42.3	934	4 Q9B2F3	Q9B2F3 homo sapien
19	43	41.3	113	2 Q9K5A5	Q9K5A5 enterobacte
20	43	41.3	167	16 Q99WL1	Q99WL1 staphylococ
21	43	41.3	305	16 Q31A43	Q31A43 bacillus su
22	43	41.3	432	6 Q95LX5	Q95LX5 macaca fasc
23	43	41.3	445	11 Q9D3B7	Q9D3B7 mus musculu
24	43	41.3	446	11 Q9D611	Q9D611 mus musculu
25	43	41.3	453	11 Q925E7	Q925E7 mus musculu
26	43	41.3	457	11 Q9CW40	Q9CW40 mus musculu
27	43	41.3	468	13 P87345	P87345 xenopus lae
28	43	41.3	468	13 Q92141	Q92141 xenopus lae
29	43	41.3	479	11 Q9QW44	Q9QW44 rattus sp.
30	43	41.3	479	11 Q9QWS2	Q9QWS2 mus musculu
31	42.5	40.9	748	4 Q16149	Q16149 homo sapien
32	42	40.4	139	17 O59457	O59457 pyrococcus
33	42	40.4	142	9 Q9MBS1	Q9MBS1 staphylococ
34	42	40.4	155	5 O16597	O16597 caenorhabdi
35	42	40.4	206	16 Q9JRA2	Q9JRA2 drosophila
36	42	40.4	225	5 Q9VH21	Q9VH21 vibrio chol
37	42	40.4	393	16 Q9KTX2	Q9KTX2 vibrio chol
38	42	40.4	477	4 Q9PIY7	Q9PIY7 homo sapien
39	42	40.4	514	16 Q9CKM5	Q9CKM5 pasteurella
40	42	40.4	608	11 Q9QXK1	Q9QXK1 mus musculu
41	42	40.4	665	8 Q9G8V7	Q9G8V7 rhodomonas
42	42	40.4	851	16 Q91442	Q91442 pseudomonas
43	41	39.4	136	2 O82915	O82915 escherichia
44	41	39.4	158	17 O58392	O58392 pyrococcus
45	41	39.4	210	11 Q9CU40	Q9CU40 mus musculu

ALIGNMENTS

RESULT 1
Q99015 PRELIMINARY; PRT; 46 AA.
ID Q99015
AC Q99015
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN PRECURSOR.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB14, H7, LT11, NG8, AND UA159;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in
RT Biofilms.";
RL J. Bacteriol. 183:897-908(2001).
DR EMBL; AF277152; AAK01542.1; -
DR EMBL; AF277153; AAK01543.1; -
DR EMBL; AF277155; AAK01545.1; -
DR EMBL; AF277156; AAK01546.1; -
DR EMBL; AF277157; AAK01547.1; -
DR InterPro; IPR004288; ComC.
DR Pfam; PF03047; ComC; 1.
FT CHAIN 26 46
SQ SEQUENCE 46 AA; 5211 MW; 38FA62B6F78FC3BF CRC64;
Query Match 100.0%; Score 104; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLNRSFTQALGK 21

Db 26 SGSLSTFFRLNRSFTQALGK 46

```

RESULT 2
Q9APK7 PRELIMINARY; PRT; 46 AA.
ID Q9APK7; 2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM71;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in
RT Biofilms.";
RL J. Bacteriol. 183:897-908(2001).
DR EMBL: AF277151; AAK01541.1;
FT CHAIN 26 46
SQ SEQUENCE 46 AA; 5195 MW; 38E0B9DB5B8FC3BF CRC64;

Query Match 100.0%; Score 104; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGLSTFFRLFNRSFTQALGK 21
DB 26 SGLSTFFRLFNRSFTQALGK 46

RESULT 3
Q9APK6 PRELIMINARY; PRT; 43 AA.
ID Q9APK6; 2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JHI005;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in
RT Biofilms.";
RL J. Bacteriol. 183:897-908(2001).
DR EMBL: AF277154; AAK01544.1;
FT CHAIN 26 43
SQ SEQUENCE 43 AA; 4927 MW; E6A78FC3B6156C7 CRC64;

Query Match 82.7%; Score 86; DB 2; Length 43;
Best Local Similarity 94.4%; Pred. No. 1.8e-07;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGLSTFFRLFNRSFTQA 18
DB 26 SGLSTFFRLFNRSFTQA 43

RESULT 4
Q5584 PRELIMINARY; PRT; 1083 AA.
ID Q5584; 2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE ACRYLAMIDE RESISTANCE PROTEIN.
GN ENVD OR SRO369.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC 6803;
RA Tabata S.;
RX Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nariu K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL: D63999; BAA10089.1;
DR InterPro; IPR001036; ACR_tran.
DR Pfam; PF00873; ACR_tran; I.
DR PRINTS; PR00702; ACRYLAMINRP.
KW Complete proteome.
SQ SEQUENCE 1083 AA; 117561 MW; 4388B79D6BC177A CRC64;

Query Match 46.2%; Score 48; DB 16; Length 1083;
Best Local Similarity 64.3%; Pred. No. 17;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 SGLSTFFRLFNRSF 15
DB 514 GPLAWFFNLENRTP 527

RESULT 5
Q9QZS6 PRELIMINARY; PRT; 390 AA.
ID Q9QZS6; 2001 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE D-GLYCOSAMINYL 3-O-SULFOTRANSFERASE-3B.
GN HS3T3B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=99449300; PubMed=10520990;
RA Shukla D., Liu J., Blaiklock P., Shworak N.W., Bai X., Esko J.D.,
RA Cohen G.H., Eisenberg R.J., Rosenberg R.D., Spear P.G.;
RT "A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus
RT 1 entry.";
RL Cell 99:13-22(1999).
DR EMBL: AF168992; AAF04505.1;

```

```
DR MGD: MGI:1333853; Hs3st3b.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.
KW Transferase.
SQ SEQUENCE 390 AA; 43326 MW; ACD0D28D6B3DDE8 CRC64;

Query Match 43.3%; Score 45; DB 11; Length 390;
Best Local Similarity 50.0%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY 4 LSTFFRLNRSFTQALGK 21
| | | | | | | | | |
Db 368 LRDYFDPENRKFYQMTGR 385

RESULT 6
Q9QNB1 ID Q9QNB1 PRELIMINARY; PRT; 835 AA.
AC Q9QNB1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VP3.
GN VP3.
OS Human rotavirus (strain KU).
OC Viruses; dsRNA viruses; Reoviridae; Rotavirus.
OX NCBI_TaxID=10952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KU;
RA Taniguchi K.;
RT "Rotavirus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB022767; BAA84964.1; -.
SQ SEQUENCE 835 AA; 97857 MW; EA5D8227C147E599 CRC64;

Query Match 43.3%; Score 45; DB 12; Length 835;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLFN 12
| | | | | | | |
Db 768 SGVLSYFKLYN 779

RESULT 7
Q97N40 ID Q97N40 PRELIMINARY; PRT; 850 AA.
AC Q97N40;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE ABC TRANSPORTER, PERMEASE PROTEIN, PUTATIVE.
GN SP2231.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
RT pneumoniae.";
RL Science 293:498-506(2001).
```

```
DR EMBL: AE007510; AAK76279.1; -.
DR TIGR: SP2231; -.
KW Complete proteome.
SQ SEQUENCE 850 AA; 97303 MW; 1ADE613F06B5115 CRC64;

Query Match 43.3%; Score 45; DB 16; Length 850;
Best Local Similarity 50.0%; Pred. No. 43;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 6 TFRFLNRSFTQALGK 21
| | | | | | | | | |
Db 733 TFRFLNRSFTQALGK 748

RESULT 8
Q991N1 ID Q991N1 PRELIMINARY; PRT; 1528 AA.
AC Q991N1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)
DE HELICASE (FRAGMENT).
OS Little cherry virus-2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus.
OX NCBI_TaxID=154339;
RN [1]
RP SEQUENCE FROM N.A.
RA Rott M.E., Jelkmann W.;
RT "Identification of a second closterovirus associated with little
RT cherry disease, little cherry virus-2.";
RL Phytopathology 91:0-0(2001).
DR EMBL: AF33237; AAK19543.1; -.
DR InterPro: IPR000606; Viral_helicase.
DR InterPro: IPR002588; V_methyltransf.
DR Pfam: PF01443; Viral_helicase; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
FT NON_TER 1
SQ SEQUENCE 1528 AA; 170263 MW; 1AA54A7016AE27E2 CRC64;

Query Match 43.3%; Score 45; DB 12; Length 1528;
Best Local Similarity 45.0%; Pred. No. 80;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 GSLSSTFFRLNRSFTQALGK 21
| | | | | | | | | |
Db 325 GSLSSTFFRLNRSFTQALGK 344

RESULT 9
Q991N2 ID Q991N2 PRELIMINARY; PRT; 1980 AA.
AC Q991N2;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE RNA DEPENDENT RNA POLYMERASE (FRAGMENT).
OS Little cherry virus-2.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Closteroviridae;
OC Closterovirus.
OX NCBI_TaxID=154339;
RN [1]
RP SEQUENCE FROM N.A.
RA Rott M.E., Jelkmann W.;
RT "Identification of a second closterovirus associated with little
RT cherry disease, little cherry virus-2.";
RL Phytopathology 91:0-0(2001).
DR EMBL: AF33237; AAK19544.1; -.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase.
DR InterPro: IPR002588; V_methyltransf.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase; 1.
```


Matches	9;	Conservative	2;	Mismatches	7;	Indels	0;	Gaps	0;
---------	----	--------------	----	------------	----	--------	----	------	----

QY	4	LSTFFRLFNRSFTQALGK 21
		: :
Db	359	LOEYRPFNRRFYQMTCG 376


```

RESULT 12
ID Q9N3H3 PRELIMINARY; PRT; 431 AA.
AC Q9N3H3;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2001 (TremBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE HYPOTHETICAL 48.3 KDA PROTEIN.
GN Y53G8AL.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN RNN
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium." ;
RL Science 282:2012-2018(1998).
[2]
RN RNN
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Latreille P.;
RT "The sequence of C. elegans cosmid Y53G8AL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
[3]
RN RNN
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AC024807; AAF59528.2; --
KW Hypothetical protein.
SQ SEQUENCE 431 AA; 48254 MW; D5ADA2DDC3952A2C CRC64;
Query Match 42.3%; Score 44; DB 5; Length 431;
Best Local Similarity 38.1%; Pred. No. 30;
Matches 8; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

```


QY	1	SGSLTFPFLFNRSFTQALGK 21
	:	: :
Db	375	AGGQQAYRFSNRYFEEQYE 395


```

RESULT 13
ID Q96SR1 PRELIMINARY; PRT; 719 AA.
AC Q96SR1;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE CDNA FLJ114694 FIS, CLONE NT2RP2005407, WEAKLY SIMILAR TO
DE OXYSTEROL-BINDING PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RNN
RP SEQUENCE FROM N.A.
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA Matsunaga H., Ishii S., Kawai I., Saito K., Yamamoto J., Wakamatsu A.,
RA Nakamura Y., Nagahari K., Masuo Y., Sasaki N.;
RT "NEO human cDNA sequencing project.";

```


RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK07600; BAB5223.1; -.
 SQ SEQUENCE 719 AA; 81876 MW; A2B10D14D265A41E CRC64;

Query Match 42.3%; Score 44; DB 4; Length 719;
 Best Local Similarity 42.9%; Pred No. 53;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 SGSLSTFFFLNRSFTQALGK 21
 || |||| : : ||
 Db 419 SGYCSTYFRAGSKPFPVLGE 439

RESULT 14

OY9N007 PRELIMINARY; PRT; 740 AA.

AC OY9N007;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 84.2 KDA PROTEIN.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-CEREBELLUM CORTEX;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.;

RT "Isolation of full-length cDNA clones from macaque brain cDNA

RT libraries."

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB046640; BAB03558.1; -.

DR InterPro: IPR000648; Oxysterol_BP.

DR Pfam; PF01237; Oxysterol_BP; 1.

DR PROSITE; PS01013; OSBP; 1.

KW Hypothetical protein.

SQ SEQUENCE 740 AA; 84223 MW; 6B8E850FD9E13C78 CRC64;

Query Match 42.3%; Score 44; DB 6; Length 740;
 Best Local Similarity 42.9%; Pred No. 54;
 Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 SGSLSTFFFLNRSFTQALGK 21
 || |||| : : ||
 Db 440 SGYCSTYFRAGSKPFPVLGE 460

RESULT 15

OY9DB10 PRELIMINARY; PRT; 799 AA.

AC OY9DB10;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE 1300008A22RIK PROTEIN.
 GN 1300008A22RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=LIVER;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata Y., Konno H., Adachi J., Fukuda S.,

RA Atakawa T., Hara A., Fukushima Y., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner N., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 11; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 CC EMBL; AK004939; BAB23684.1; -.
 DR HSP; P00763; IDPO.
 DR MGI; MGI:1919003; 1300008A22RIK.
 DR InterPro: IPR001314; Chymotrypsin.
 DR InterPro: IPR000859; CUB.
 DR InterPro: IPR002172; LDL_recept_A.
 DR InterPro: IPR001254; trypsin.
 DR Pfam; PF00057; ldl_recept_a; 3.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00192; LDLa; 3.
 DR SMART; SM00020; Tryp_Spc; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS50068; LDLRA_2; 3.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN.1.
 KW Glycoprotein; Hydrolase; Serine protease.
 SQ SEQUENCE 799 AA; 89557 MW; 16315A646AD5288 CRC64;

Query Match 42.3%; Score 44; DB 11; Length 799;
 Best Local Similarity 52.4%; Pred No. 59;
 Matches 11; Conservative 3; Mismatches 3; Indels 4; Gaps 1;

OY 1 SGSLSTFFFLNRSFTQALGK 21
 |||| : : ||||
 Db 79 SGSL----RVLNRHFSQDLGR 95

Search completed: November 5, 2002, 10:57:35
 Job time : 18.4776 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:57:42 ; Search time 16.6119 Seconds
(without alignments)
140.414 Million cell updates/sec

Title: US-09-833-017B-4
Perfect score: 21
Sequence: 1 SGSLSTFFRLNRSFTQALCK 21

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 11073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : A_Geneseq_032802.*

1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7	33.3	929	22 ABB70857	Drosophila melanog
2	6	28.6	27	20 AAY40080	Peptide sequence d
3	6	28.6	48	22 AAM23675	Human EST encoded
4	6	28.6	51	17 AAR94979	Calcium channel is
5	6	28.6	51	17 AAR94978	Calcium channel is
6	6	28.6	62	17 AAR97851	Rat brain calcium
7	6	28.6	66	17 AAR97852	Rat brain calcium
8	6	28.6	66	22 ABG04703	Novel human diago
9	6	28.6	78	22 AAM93377	Human immune/haema
10	6	28.6	96	19 AAM93796	Nucleotide sequenc
11	6	28.6	122	20 AAY40074	Peptide sequence d

12	6	28.6	141	22 AAU45360	Propionibacterium
13	6	28.6	147	22 AAE11946	Human CCAAT/enhanc
14	6	28.6	150	22 AAU16099	Human novel secret
15	6	28.6	152	21 AAG48562	Arabidopsis thalia
16	6	28.6	180	19 AAG59797	Amino acid sequenc
17	6	28.6	186	21 AAG19904	Arabidopsis thalia
18	6	28.6	194	22 AAU19473	Human diagnostic a
19	6	28.6	212	22 ABG24830	Chlamydia trachoma
20	6	28.6	222	20 AAY37757	Novel human diago
21	6	28.6	254	22 ABG12787	Novel human diago
22	6	28.6	254	22 ABG15795	Novel human diago
23	6	28.6	257	21 AAG19903	Arabidopsis thalia
24	6	28.6	270	13 AAR27641	Human calcium chan
25	6	28.6	279	22 ABB16367	Human nervous syst
26	6	28.6	279	22 AAU16524	Human novel secret
27	6	28.6	282	22 AAU05262	Chicken stem cell
28	6	28.6	282	22 AAB98364	Chicken SCF protei
29	6	28.6	282	22 AAU02489	Chicken SCF (stem
30	6	28.6	282	22 AAU02775	Chicken SCF (stem
31	6	28.6	282	22 AAB73576	Chicken SCF (stem
32	6	28.6	282	22 AAB96949	Chicken stem cell
33	6	28.6	286	22 AAU32713	Novel human secret
34	6	28.6	287	21 AAG19902	Arabidopsis thalia
35	6	28.6	296	22 AAE11947	Mouse CCAAT/enhanc
36	6	28.6	297	22 AAE11948	Rat CCAAT/enhanc
37	6	28.6	322	22 AAE11945	Human CCAAT/enhanc
38	6	28.6	329	22 AAU35549	Haemophilus influe
39	6	28.6	329	22 ABG11238	Novel human diago
40	6	28.6	334	22 AAU57978	Propionibacterium
41	6	28.6	344	22 ABG03199	Novel human diago
42	6	28.6	345	12 AAR14408	Nuclear factor C/E
43	6	28.6	345	22 AAE11944	Human CCAAT/enhanc
44	6	28.6	345	22 AAU09068	Human transcriptio
45	6	28.6	356	22 AAM39772	Human polypeptide
46	6	28.6	359	22 ABB70995	Drosophila melanog
47	6	28.6	360	22 AAY34924	Chlamydia pneumoni
48	6	28.6	360	21 AAY59364	Amino acid sequenc
49	6	28.6	360	22 ABG14697	Novel human diago
50	6	28.6	360	22 AAM41558	Human polypeptide
51	6	28.6	372	22 AAU48589	Propionibacterium
52	6	28.6	398	22 ABG10841	Novel human diago
53	6	28.6	412	21 AAG48540	Arabidopsis thalia
54	6	28.6	414	21 AAG48539	Arabidopsis thalia
55	6	28.6	446	20 AAY30792	Mature secreted hum
56	6	28.6	447	21 AAG48538	Arabidopsis thalia
57	6	28.6	465	22 ABG03158	Novel human diago
58	6	28.6	466	20 AAY30793	Immature human HKN
59	6	28.6	477	20 AAY30786	Protein encoded by
60	6	28.6	477	22 AAB94726	Human protein sequ
61	6	28.6	478	22 ABB58996	Drosophila melanog
62	6	28.6	489	13 AAR27647	Human calcium chan
63	6	28.6	495	20 AAY30785	Protein encoded by
64	6	28.6	499	13 AAR27574	The ABF-B from A.
65	6	28.6	513	22 AAM41685	Human polypeptide
66	6	28.6	518	13 AAR27644	Human calcium chan
67	6	28.6	559	21 AAY91945	Human chapetone pr
68	6	28.6	559	22 AAB93115	Human protein sequ
69	6	28.6	561	22 AAU34618	E. coli cellular p
70	6	28.6	595	22 AAB38334	Salmonella typhi c
71	6	28.6	600	21 AAG40930	Zea mays protein f
72	6	28.6	635	19 AAW75855	Human secretory pr
73	6	28.6	635	20 AAY45160	Human secreted pro
74	6	28.6	635	21 AAY71060	Human membrane tra
75	6	28.6	684	13 AAR27645	Human calcium chan
76	6	28.6	698	21 AAY99666	Human GTPase assoc
77	6	28.6	703	15 AAR49135	Sequence of lag D
78	6	28.6	908	22 AAM00834	Human bone marrow
79	6	28.6	993	13 AAR27650	Human calcium chan
80	6	28.6	1087	22 ABB24996	Novel human diago
81	6	28.6	1127	22 ABG22813	Novel human diago
82	6	28.6	1217	22 AAU33177	Novel human secret
83	6	28.6	1343	22 AAM00947	Human bone marrow
84	6	28.6	1359	21 AAB41785	Human OREF ORF1549

85	6	28.6	1820	22	AAU27679	Human full-length	158	5	23.8	44	22	ABG28609	Novel human diago
86	6	28.6	1824	21	AAU49431	Murine CACNAIF pro	159	5	23.8	46	20	AAU30849	Human secreted pro
87	6	28.6	1912	21	AAU49429	Human CACNAIF spli	160	5	23.8	46	20	AAU89492	Human immune/haema
88	6	28.6	1967	14	AAU33547	Sequence of the al	161	5	23.8	46	22	AAU76210	Human colon cancer
89	6	28.6	1968	19	AAU63139	Human calcium chan	162	5	23.8	47	22	ABU1431	Peptide #8937 enco
90	6	28.6	1977	21	AAU49430	Human CACNAIF long	163	5	23.8	47	22	AAU62304	Human brain expres
91	6	28.6	2138	16	AAU72607	Human neuronal cal	164	5	23.8	47	22	AAU75107	Human bone marrow
92	6	28.6	2138	21	AAU10593	Human neuronal cal	165	5	23.8	47	22	AAU35223	Peptide #9260 enco
93	6	28.6	2161	14	AAU33545	Human calcium chan	166	5	23.8	48	22	AAU76454	Human colon cancer
94	6	28.6	2161	16	AAU71001	Sequence of the al	167	5	23.8	49	22	ABU42096	Peptide #9602 enco
95	6	28.6	2161	16	AAU71002	Human neuronal cal	168	5	23.8	49	22	AAU62976	Human brain expres
96	6	28.6	2161	19	AAU63137	Human calcium chan	169	5	23.8	49	22	AAU75789	Human bone marrow
97	6	28.6	2161	19	AAU63149	Human calcium chan	170	5	23.8	49	22	AAU35899	Peptide #9936 enco
98	6	28.6	2161	21	AAU63156	Human calcium chan	171	5	23.8	50	22	AAU50306	Propionibacterium
99	6	28.6	2163	16	AAU71003	Human calcium chan	172	5	23.8	51	17	AAU94977	Calcium channel is
100	6	28.6	2163	21	AAU10570	Human neuronal cal	173	5	23.8	52	21	AAU00521	Human secreted pro
101	6	28.6	2279	22	AAU38753	Human polypeptide	174	5	23.8	52	22	ABU16203	Human nervous syst
102	6	28.6	2297	22	AAU40519	Human polypeptide	175	5	23.8	52	22	AAU06584	Human foetal prote
103	6	28.6	2322	18	AAU15566	Melanoma-associate	176	5	23.8	52	22	AAU06729	Human secreted pro
104	6	28.6	2344	22	AAU37120	Staphylococcus aur	177	5	23.8	54	21	AAU38489	Human nervous syst
105	5	23.8	5	14	AAU44624	Calcitonin fragmen	178	5	23.8	54	22	ABU17231	Peptide #2128 enco
106	5	23.8	8	19	AAU13431	Melanoma-homing pe	179	5	23.8	55	22	ABU29477	Peptide #3651 enco
107	5	23.8	8	19	AAU60292	Melanoma-tumour ho	180	5	23.8	55	22	ABU31000	Peptide #2161 enco
108	5	23.8	8	20	AAU931716	Mouse B16B15b mela	181	5	23.8	55	22	ABU34655	Peptide #3699 enco
109	5	23.8	8	21	AAU21704	Murine melanoma ho	182	5	23.8	55	22	ABU36193	Human nervous syst
110	5	23.8	8	21	AAU33539	Angiogenic vascula	183	5	23.8	55	22	ABU15421	Human nervous syst
111	5	23.8	8	22	AAU79611	Integrin-binding p	184	5	23.8	55	22	ABU20062	Protein #2061 enco
112	5	23.8	8	22	AAU06282	Tumour homing pept	185	5	23.8	55	22	AAU55445	Human brain expres
113	5	23.8	13	14	AAU40791	Sequence #4 from a	186	5	23.8	55	22	AAU56974	Human brain expres
114	5	23.8	15	20	AAU22379	TPO receptor bindi	187	5	23.8	55	22	AAU60182	Human bone marrow
115	5	23.8	15	20	AAU02605	Mouse beta-actin a	188	5	23.8	55	22	AAU67836	Human bone marrow
116	5	23.8	15	20	AAU92534	Beta-actin referen	189	5	23.8	55	22	AAU69361	Human bone marrow
117	5	23.8	16	15	AAU7026	Plasminogen activa	190	5	23.8	55	22	AAU72800	Human bone marrow
118	5	23.8	19	19	AAU69830	Peptide fragment p	191	5	23.8	55	22	AAU15651	Peptide #2085 enco
119	5	23.8	19	21	AAU85614	Cathepsin-D antige	192	5	23.8	55	22	AAU28152	Peptide #2189 enco
120	5	23.8	21	18	AAU34053	Human MDR1-P glyco	193	5	23.8	55	22	AAU33030	Peptide #7067 enco
121	5	23.8	23	16	AAU5099	Random biotinylati	194	5	23.8	55	22	AAU04888	Peptide #2072 enco
122	5	23.8	23	19	AAU46702	Biotinylation pept	195	5	23.8	55	22	AAU23726	Human EST encoded
123	5	23.8	23	20	AAU29305	Biotinylation pept	196	5	23.8	57	22	AAU24492	Human secreted pro
124	5	23.8	23	20	AAU67910	Human secreted pro	197	5	23.8	60	21	AAU02747	Propionibacterium
125	5	23.8	23	21	AAU87142	Human secreted pro	198	5	23.8	60	22	AAU56956	Novel human diago
126	5	23.8	23	21	AAU54984	Angiogenin inhibit	199	5	23.8	61	22	ABU04669	Novel human diago
127	5	23.8	23	22	AAU04311	ATP-binding casset	200	5	23.8	61	22	ABU09797	Human immune/haema
128	5	23.8	23	22	AAU06119	Human gene 14 enco	201	5	23.8	61	22	AAU83024	Human secreted pro
129	5	23.8	23	22	AAU09937	Biotinylated pepti	202	5	23.8	62	21	AAU34466	C glutamicum prote
130	5	23.8	25	20	ABU01293	Novel human diago	203	5	23.8	62	22	AAU91645	Zea mays protein f
131	5	23.8	27	20	AAU67973	Fragment of human	204	5	23.8	63	21	AAU18764	Zea mays protein f
132	5	23.8	30	22	AAU08982	Human polypeptide	205	5	23.8	63	21	AAU32669	Arabidopsis thalia
133	5	23.8	31	7	AAU61348	Sequence of calci	206	5	23.8	63	21	AAU48401	Human haematologic
134	5	23.8	31	7	AAU61349	Sequence of calci	207	5	23.8	63	22	AAU81044	Human haematologic
135	5	23.8	31	22	AAU75470	Human bone marrow	208	5	23.8	63	22	AAU82041	Human haematologic
136	5	23.8	32	17	AAU03618	Rat bradykinin GPR	209	5	23.8	63	22	AAU95167	Human reproductive
137	5	23.8	32	22	AAU92939	C glutamicum prote	210	5	23.8	65	22	AAU58313	Propionibacterium
138	5	23.8	34	22	ABU08616	Novel human diago	211	5	23.8	65	22	ABU39031	Propionibacterium
139	5	23.8	35	22	ABU15474	Peptide #5316 enco	212	5	23.8	65	22	ABU23975	Protein #5974 enco
140	5	23.8	35	22	AAU18882	Human secreted pep	213	5	23.8	65	22	AAU59685	Human brain expres
141	5	23.8	36	21	AAU38977	Human secreted pep	214	5	23.8	65	22	AAU72361	Human bone marrow
142	5	23.8	36	22	ABU10490	Novel human diago	215	5	23.8	65	22	AAU19520	Peptide #5954 enco
143	5	23.8	39	18	AAU27950	Staphylococcus aur	216	5	23.8	65	22	AAU32520	Peptide #6557 enco
144	5	23.8	40	9	AAU80076	Cytomegalovirus (C	217	5	23.8	65	22	AAU41112	Human polypeptide
145	5	23.8	40	21	AAU55145	Arabidopsis thalia	218	5	23.8	65	22	AAU41112	Human endometrium
146	5	23.8	40	22	ABU22221	Protein #4220 enco	219	5	23.8	66	20	AAU60064	Peptide #11640 enc
147	5	23.8	40	22	AAU01625	Human secreted pro	220	5	23.8	66	22	ABU44134	Protein #5021 enco
148	5	23.8	41	21	AAU50421	Human uncoupling p	221	5	23.8	66	22	ABU23022	Protein #9015 enco
149	5	23.8	41	21	AAU71164	HPIV2 partial prot	222	5	23.8	66	22	ABU27016	Human brain expres
150	5	23.8	41	22	ABU31202	Peptide #3853 enco	223	5	23.8	66	22	AAU65157	Human bone marrow
151	5	23.8	41	22	ABU36400	Peptide #3906 enco	224	5	23.8	66	22	AAU77862	Human brain expres
152	5	23.8	41	22	AAU57165	Human brain expres	225	5	23.8	66	22	AAU21762	Peptide #8196 enco
153	5	23.8	41	22	AAU09564	Human bone marrow	226	5	23.8	66	22	AAU38084	Peptide #12121 enc
154	5	23.8	41	22	AAU73385	Peptide #3819 enco	227	5	23.8	67	22	AAU94910	Human reproductive
155	5	23.8	41	22	AAU29904	Peptide #3941 enco	228	5	23.8	67	22	AAU11655	Human polypeptide
156	5	23.8	41	22	AAU05074	Peptide #3756 enco	229	5	23.8	68	20	AAU07905	Human secreted pro
157	5	23.8	43	22	AAU62905	Human TNF receptor	230	5	23.8	68	22	ABU00305	Novel human diago

231	5	23.8	68	22	AAO13010	Human polypeptide	304	5	23.8	92	22	ABG08059	Novel human diagno
232	5	23.8	69	22	ABB30649	Peptide #3300 enco	305	5	23.8	93	22	AAU56795	Propionibacterium
233	5	23.8	69	22	ABB35820	Peptide #3326 enco	306	5	23.8	93	22	ABG19225	Novel human diagno
234	5	23.8	69	22	ABB21236	Protein #3235 enco	307	5	23.8	94	19	AAW79096	Human secreted pro
235	5	23.8	69	22	AAAG69004	Human bone marrow	308	5	23.8	94	21	AAW57028	Human prostate can
236	5	23.8	69	22	AAAG90904	Human immune/haema	309	5	23.8	94	21	AAAG55575	Arabidopsis thalia
237	5	23.8	69	22	AAAG39321	Peptide #3358 enco	310	5	23.8	94	22	ABG38213	Novel human diagno
238	5	23.8	70	17	AAU97850	Rat brain calcium	311	5	23.8	95	18	AAW28022	Amino acid sequenc
239	5	23.8	70	22	AAU50030	Propionibacterium	312	5	23.8	95	20	AAU59742	Human normal ovari
240	5	23.8	70	22	AAU59698	Propionibacterium	313	5	23.8	95	21	AAAG60424	Arabidopsis thalia
241	5	23.8	71	21	AAAG38763	Zea mays protein f	314	5	23.8	96	22	AAU56723	Propionibacterium
242	5	23.8	71	21	AAAG32688	Zea mays protein f	315	5	23.8	97	12	AAU12123	TRFP chain 2 - t
243	5	23.8	71	21	AAAG37983	Arabidopsis thalia	316	5	23.8	97	18	AAU11247	Streptococcus pneu
244	5	23.8	71	21	AAAG48400	Arabidopsis thalia	317	5	23.8	97	19	AAW38634	T cell reactive fe
245	5	23.8	73	20	AAAG73882	Human prostate tum	318	5	23.8	97	21	AAW28935	Feline human TRFP
246	5	23.8	73	22	ABG37287	Novel human diagno	319	5	23.8	97	21	AAU87675	Cat TRFP chain 2 t
247	5	23.8	74	22	AAU56270	Propionibacterium	320	5	23.8	97	21	AAU91015	Human TRFP chain 2
248	5	23.8	74	22	AAU39326	Human polypeptide	321	5	23.8	97	21	AAU51472	Propionibacterium
249	5	23.8	75	21	AAU08492	Amino acid sequenc	322	5	23.8	97	22	AAU45726	Propionibacterium
250	5	23.8	75	22	AAAG67855	Murine leukemia vi	323	5	23.8	97	22	AAU46334	Propionibacterium
251	5	23.8	76	22	AAAB4253	Human immune/haema	324	5	23.8	97	22	AAU10505	Propionibacterium
252	5	23.8	77	21	AAAB34363	Human secreted pro	325	5	23.8	98	22	AAU67006	Propionibacterium
253	5	23.8	77	22	ABB28103	Human peptide #754	326	5	23.8	98	22	AAU42269	Propionibacterium
254	5	23.8	77	22	ABB33278	Peptide #784 enco	327	5	23.8	98	22	AAU55100	Propionibacterium
255	5	23.8	77	22	ABB32780	Protein #739 enco	328	5	23.8	100	21	AAU10261	Arabidopsis thalia
256	5	23.8	77	22	ABB18740	Human brain expres	329	5	23.8	100	22	AAAB1404	Human haematologic
257	5	23.8	77	22	AAU54070	Human bone marrow	330	5	23.8	100	22	AAAB1907	Human haematologic
258	5	23.8	77	22	AAU66458	Human bone marrow	331	5	23.8	101	13	AAU27372	TRFP Chain #2 C2ST
259	5	23.8	77	22	AAU14330	Peptide #764 enco	332	5	23.8	101	13	AAU27372	Human prostate can
260	5	23.8	77	22	AAU26740	Peptide #777 enco	333	5	23.8	101	21	AAU57154	Human ORFX ORF2040
261	5	23.8	78	18	AAU20059	Peptide #741 enco	334	5	23.8	101	21	AAU42276	Propionibacterium
262	5	23.8	78	18	AAU23001	Canine herpesvirus	335	5	23.8	101	22	AAU40192	Propionibacterium
263	5	23.8	78	19	AAU27261	Canine herpes viru	336	5	23.8	102	22	AAU66242	Human immune/haema
264	5	23.8	78	20	AAU08239	Human	337	5	23.8	102	22	AAU86328	Human immune/haema
265	5	23.8	78	22	AAU2026	Human haematologic	338	5	23.8	103	20	AAU74435	Human colon cancer
266	5	23.8	78	22	AAU64235	Human brain expres	339	5	23.8	103	20	AAU12429	Human 5' EST secre
267	5	23.8	79	21	AAU51318	Canine herpes viru	340	5	23.8	103	21	AAU40704	Human ORFX ORF468
268	5	23.8	79	21	AAU3069	Human ORFX ORF2833	341	5	23.8	103	22	AAAB1878	Human haematologic
269	5	23.8	80	22	AAU04436	Arabidopsis thalia	342	5	23.8	104	15	AAAB1942	Human haematologic
270	5	23.8	81	22	AAU23267	Novel human enzyme	343	5	23.8	104	15	AAAB5453	T-cell receptor V-
271	5	23.8	82	22	AAU42097	Propionibacterium	344	5	23.8	105	20	AAAG24831	Arabidopsis thalia
272	5	23.8	82	22	AAU67375	Propionibacterium	345	5	23.8	105	20	AAU37921	Chlamydia trachoma
273	5	23.8	82	22	AAU27371	TRFP Chain #2 C2ST	346	5	23.8	105	22	AAAG6359	Human brain expres
274	5	23.8	82	22	AAU23486	Human EST encoded	347	5	23.8	105	22	AAAG76402	Human bone marrow
275	5	23.8	84	22	AAAB9076	Human secreted pro	348	5	23.8	105	22	AAU36511	Peptide #10548 enc
276	5	23.8	84	22	AAU06597	Human foetal prote	349	5	23.8	106	22	AAAG62631	C Sorokiniana omeg
277	5	23.8	84	22	AAU06793	Human foetal prote	350	5	23.8	106	12	AAU15449	Asn 9 pGRF precurs
278	5	23.8	85	21	AAAG37314	Arabidopsis thalia	351	5	23.8	106	12	AAU15449	Ser 9 pGRF precurs
279	5	23.8	85	21	AAAG48399	Arabidopsis thalia	352	5	23.8	106	22	AAU23520	Amino acid sequenc
280	5	23.8	85	22	AAU07257	Human polypeptide	353	5	23.8	106	22	AAU50536	Propionibacterium
281	5	23.8	86	21	AAU50420	Human octapeptid p	354	5	23.8	106	22	AAAB7452	Protein acid sequenc
282	5	23.8	86	22	AAU42401	Propionibacterium	355	5	23.8	106	22	AAAB7453	Fab' Ch1. Synthet
283	5	23.8	86	22	AAU62283	Propionibacterium	356	5	23.8	106	22	AAAG63351	Fab' H-N162D. Syn
284	5	23.8	87	22	ABG08617	Novel human diagno	357	5	23.8	106	22	AAAG63352	Amino acid sequenc
285	5	23.8	87	21	AAU56339	Arabidopsis thalia	358	5	23.8	107	21	AAAB18440	Human ORFX ORF1604
286	5	23.8	88	20	AAU73898	Human prostate tum	359	5	23.8	107	21	AAAG26469	Arabidopsis thalia
287	5	23.8	88	21	AAU37313	Arabidopsis thalia	360	5	23.8	107	21	AAAG47776	Arabidopsis thalia
288	5	23.8	88	22	AAU21374	Human novel foetal	361	5	23.8	107	21	AAAG48382	Arabidopsis thalia
289	5	23.8	88	22	AAU06338	Human foetal prote	362	5	23.8	108	21	AAAG44113	Arabidopsis thalia
290	5	23.8	88	22	AAAB61901	Mouse apoptosis me	363	5	23.8	109	12	AAAR12122	TRFP I chain 2 - s
291	5	23.8	88	22	AAU29338	Human secreted pro	364	5	23.8	109	14	AAAC36541	TRFP chain 2 (with
292	5	23.8	89	21	AAU29338	Human secreted pro	365	5	23.8	109	14	AAAB1985	Human T cell react
293	5	23.8	89	21	AAU36152	Arabidopsis thalia	366	5	23.8	109	20	AAU25593	Felis sp. allergen
294	5	23.8	89	22	AAU39089	Human polypeptide	367	5	23.8	109	21	AAAB28934	Human T cell react
295	5	23.8	90	21	AAU011912	Human secreted pro	368	5	23.8	109	21	AAU87674	Feline human TRFP
296	5	23.8	90	21	AAU02990	Human secreted pro	369	5	23.8	109	21	AAU90104	Cat TRFP chain 2 s
297	5	23.8	90	22	AAU14264	Human novel protei	370	5	23.8	109	21	AAU51471	Human TRFP chain 2
298	5	23.8	90	22	AAU14500	Human novel protei	371	5	23.8	109	22	AAU20224	B. thuringiensis t
299	5	23.8	91	21	AAU18762	Zea mays protein f	372	5	23.8	110	13	AAU27370	TRFP Chain #2 with
300	5	23.8	91	21	AAAG37981	Arabidopsis thalia	373	5	23.8	110	20	AAAG37848	Chlamydia trachoma
301	5	23.8	91	22	ABG21730	Novel human diagno	374	5	23.8	111	21	AAAG47775	Arabidopsis thalia
302	5	23.8	91	22	AAAB5962	Human immune/haema	375	5	23.8	111	12	AAAR12121	TRFP chain 2 - lon
303	5	23.8	92	15	AAU65454	T-cell receptor V-	376	5	23.8	111	13	AAAB27369	Human ORFX ORF2247
304	5	23.8	92	20	AAU37880	Chlamydia trachoma	377	5	23.8	111	21	AAAB42483	Human ORFX ORF2247

377	5	23.8	111	21	AAB88933	T cell reactive fe	450	5	23.8	129	22	AAM81819	Human haematologic
378	5	23.8	111	21	AAW7673	Feline human TRFP	451	5	23.8	130	19	AAW61553	Hyalophora moth gl
379	5	23.8	111	21	AAW90103	Cat TRFP chain 2 1	452	5	23.8	130	21	AAW61553	Human cancer assoc
380	5	23.8	111	21	AAW51470	Human TRFP chain 2	453	5	23.8	130	21	AAW61553	Giant silk moth an
381	5	23.8	111	22	AAW06213	Human polypeptide	454	5	23.8	131	22	ABG29896	Novel human diagno
382	5	23.8	112	22	AAU43924	Propionibacterium	455	5	23.8	131	22	ABG29896	C glutamicum prote
383	5	23.8	112	22	AAU45902	Propionibacterium	456	5	23.8	133	22	ABG11450	Human TCFEB homolo
384	5	23.8	113	22	AAO10210	Human polypeptide	457	5	23.8	133	22	ABG11450	Corynebacterium gl
385	5	23.8	113	22	AAU02025	B. thuringiensis t	458	5	23.8	134	22	ABG11450	Drosophila melanog
386	5	23.8	114	22	AAU01082	Propionibacterium	459	5	23.8	135	22	AAU01082	Novel human secret
387	5	23.8	116	16	AAU03005	Sequence of polype	460	5	23.8	136	22	AAU03005	Pseudomonas stutze
388	5	23.8	116	16	AAU03024	Interleukin-1beta	461	5	23.8	138	21	AAU03024	Human secreted pro
389	5	23.8	116	22	AAU06442	Propionibacterium	462	5	23.8	141	22	ABG20761	Novel human diagno
390	5	23.8	116	22	AAU06517	Human des-Gln14-gh	463	5	23.8	143	20	AAU06517	H. tuberosus lecti
391	5	23.8	117	20	AAW7991	Protein designated	464	5	23.8	145	21	AAU06517	Arabidopsis thalia
392	5	23.8	117	20	AAW8008	Arabidopsis thalia	465	5	23.8	145	21	AAU06517	Arabidopsis thalia
393	5	23.8	117	21	AAW84441	Arabidopsis thalia	466	5	23.8	145	22	AAU06517	Human polypeptide
394	5	23.8	117	21	AAW84756	Arabidopsis thalia	467	5	23.8	145	22	AAU06517	Human colon cancer
395	5	23.8	117	21	AAW87236	Human signal pepti	468	5	23.8	145	22	AAU06517	Corynebacterium gl
396	5	23.8	117	22	AAO10712	Human polypeptide	469	5	23.8	146	21	AAU06517	Arabidopsis thalia
397	5	23.8	117	22	AAW38890	Human polypeptide	470	5	23.8	146	22	AAU06517	C glutamicum prote
398	5	23.8	117	22	AAW02649	Human zsig33 polyp	471	5	23.8	147	22	AAW02649	Human haematologic
399	5	23.8	117	22	AAW20101	Human zsig33 protein. H	472	5	23.8	147	22	AAW02649	Human haematologic
400	5	23.8	117	22	AAW06511	Membrane-bound pro	473	5	23.8	147	22	AAW06511	Human haematologic
401	5	23.8	118	21	AAW56708	Human ghrelin prep	474	5	23.8	147	22	AAW06511	Human secreted pro
402	5	23.8	118	22	AAW51702	Human tubulin fold	475	5	23.8	149	22	AAW51702	Novel human diagno
403	5	23.8	118	22	AAW30705	Novel human secret	476	5	23.8	149	22	AAW30705	Human polypeptide
404	5	23.8	118	22	AAO01659	Human polypeptide	477	5	23.8	149	22	AAO01659	Human secreted pro
405	5	23.8	118	22	AAU16396	Human novel secret	478	5	23.8	150	21	AAU16396	Human reproductive
406	5	23.8	118	22	AAU12392	Human PRO1066 poly	479	5	23.8	150	22	AAU12392	Mouse novel secret
407	5	23.8	118	22	AAW77380	Human colon cancer	480	5	23.8	151	19	AAW77380	Murine adult splee
408	5	23.8	118	22	AAW65231	Human PRO1066 (UNQ	481	5	23.8	151	21	AAW65231	Arabidopsis thalia
409	5	23.8	119	21	AAW18007	Arabidopsis thalia	482	5	23.8	151	21	AAW18007	Arabidopsis thalia
410	5	23.8	119	21	AAW181204	Human haematologic	483	5	23.8	151	21	AAW181204	Arabidopsis thalia
411	5	23.8	120	17	AAW62045	T. ovis vaccine ca	484	5	23.8	151	21	AAW62045	Human protein sequ
412	5	23.8	120	15	AAW92832	Taenia ovis anti	485	5	23.8	151	22	AAW92832	Human polypeptide
413	5	23.8	120	22	ABG27167	Novel human diagno	486	5	23.8	152	22	ABG27167	Novel human diagno
414	5	23.8	120	22	AAW94599	Human reproductive	487	5	23.8	152	22	AAW94599	Novel human diagno
415	5	23.8	120	22	AAO01129	Human polypeptide	488	5	23.8	152	22	AAO01129	C glutamicum prote
416	5	23.8	120	22	AAO01846	Human polypeptide	489	5	23.8	152	22	AAO01846	Propionibacterium
417	5	23.8	120	22	AAO10647	Human polypeptide	490	5	23.8	154	22	AAO10647	Novel human diagno
418	5	23.8	121	22	AAU50432	Propionibacterium	491	5	23.8	154	22	AAU50432	Novel human diagno
419	5	23.8	121	22	AAW35530	Human EST encoded	492	5	23.8	155	21	AAW35530	Human ORFX ORF508
420	5	23.8	122	15	AAW54101	Humanized MAB H-ch	493	5	23.8	156	21	AAW54101	Spinach mature lum
421	5	23.8	123	15	AAW56651	Anti-CMV monoclonal	494	5	23.8	156	22	AAW56651	Propionibacterium
422	5	23.8	123	22	ABW70162	Drosophila melanog	495	5	23.8	156	22	ABW70162	Retinol-binding pr
423	5	23.8	123	22	ABG12707	Novel human diagno	496	5	23.8	157	22	ABG12707	Spinach lumazine s
424	5	23.8	123	22	ABW37562	Human peptide #213	497	5	23.8	157	22	ABW37562	Human G protein-co
425	5	23.8	123	22	ABW32721	Peptide #227 encod	498	5	23.8	158	22	ABW32721	Novel human diagno
426	5	23.8	123	22	ABW18207	Protein #206 encod	499	5	23.8	158	22	ABW18207	Human colon cancer
427	5	23.8	123	22	ABW12232	Human secreted pro	500	5	23.8	159	22	ABW12232	Propionibacterium
428	5	23.8	123	22	AAW35539	Human brain expres	501	5	23.8	160	21	AAW35539	Protein encoded by
429	5	23.8	123	22	AAW65920	Human bone marrow	502	5	23.8	160	22	AAW65920	Drosophila melanog
430	5	23.8	123	22	AAW13787	Peptide #221 encod	503	5	23.8	160	22	AAW13787	Human reproductive
431	5	23.8	123	22	AAW26185	Peptide #222 encod	504	5	23.8	162	21	AAW26185	Murine N-terminal
432	5	23.8	123	22	AAW01532	Peptide #214 encod	505	5	23.8	162	21	AAW01532	Protein encoded by
433	5	23.8	125	19	AAW60168	Light chain variab	506	5	23.8	162	22	AAW60168	Novel human diagno
434	5	23.8	125	22	AAW31890	Novel human secret	507	5	23.8	162	22	AAW31890	Novel human diagno
435	5	23.8	126	19	AAW60167	Light chain variab	508	5	23.8	163	22	AAW60167	Propionibacterium
436	5	23.8	126	21	AAW41342	Human ORFX ORF1106	509	5	23.8	164	21	AAW41342	Arabidopsis thalia
437	5	23.8	126	21	AAW25791	Human secreted pro	510	5	23.8	164	22	AAW25791	Human AFP protein
438	5	23.8	126	21	AAW65398	Human 5' EST relat	511	5	23.8	165	22	AAW65398	Propionibacterium
439	5	23.8	126	22	ABW25140	Novel human diagno	512	5	23.8	165	22	ABW25140	Human colon cancer
440	5	23.8	126	22	AAW82236	Human immune/haema	513	5	23.8	166	21	AAW82236	Arabidopsis thalia
441	5	23.8	126	22	AAO00851	Human polypeptide	514	5	23.8	166	21	AAO00851	Arabidopsis thalia
442	5	23.8	126	22	AAW40676	Human polypeptide	515	5	23.8	166	22	AAW40676	Drosophila melanog
443	5	23.8	126	22	AAW75377	Human secreted pro	516	5	23.8	166	22	AAW75377	Paenibacillus babu
444	5	23.8	127	22	AAU50194	Propionibacterium	517	5	23.8	167	21	AAU50194	A bone marrow secr
445	5	23.8	128	22	AAU50194	Propionibacterium	518	5	23.8	167	22	AAU50194	Novel human diagno
446	5	23.8	128	22	AAU45066	Propionibacterium	519	5	23.8	167	22	AAU45066	Tomato phosphoenol
447	5	23.8	129	21	AAW41531	Arabidopsis thalia	520	5	23.8	167	22	AAW41531	Cotton phosphoenol
448	5	23.8	129	21	AAW46132	Arabidopsis thalia	521	5	23.8	168	21	AAW46132	Arabidopsis thalia
449	5	23.8	129	22	AAU57573	Propionibacterium	522	5	23.8	168	21	AAU57573	Arabidopsis thalia

523	5	23.8	168	21	AAG36151	Arabidopsis thalia	596	200	18	AAW55524	H. pylori ORF 29ep
524	5	23.8	168	21	AAG37199	Arabidopsis thalia	597	200	21	AAG47971	Arabidopsis thalia
525	5	23.8	168	21	AAG38540	Arabidopsis thalia	598	201	21	AAG11002	Arabidopsis thalia
526	5	23.8	170	19	AAG64200	Human calcineurin.	599	201	22	AAU19291	Human G protein-co
527	5	23.8	170	21	AAU09977	Human CNBII protein	600	202	14	AAW52020	Truncated human In
528	5	23.8	170	21	AAU09978	Human HCNB protein	601	202	21	AAG05783	Arabidopsis thalia
529	5	23.8	170	21	AAG09252	Arabidopsis thalia	602	203	21	AAG49184	Arabidopsis thalia
530	5	23.8	170	21	AAG09996	Arabidopsis thalia	603	204	16	AAW59481	Prolactin peptide
531	5	23.8	170	21	AAG37198	Arabidopsis thalia	604	204	21	AAG22337	Arabidopsis thalia
532	5	23.8	170	21	AAG38539	Arabidopsis thalia	605	204	21	AAW91396	Human secreted pro
533	5	23.8	172	18	AAW26532	Trypanosoma cruzi	606	205	21	AAW26660	Porphyritin antibody
534	5	23.8	172	20	AAW23294	Trypanosoma cruzi	607	205	21	AAW25288	Eucalyptus grandis
535	5	23.8	172	22	AAW23564	Novel human enzyme	608	205	22	AAW25288	Mouse apoptosis me
536	5	23.8	173	22	AAW23564	Amino acid sequenc	609	205	22	AAW37048	Arabidopsis thalia
537	5	23.8	173	22	AAW23564	Paenibacillus pabu	610	208	21	AAG47969	Arabidopsis thalia
538	5	23.8	176	21	AAG36150	Arabidopsis thalia	611	209	21	AAG47969	Arabidopsis thalia
539	5	23.8	176	22	ABW50780	Human secreted pro	612	210	22	ABW66738	Drosophila melanog
540	5	23.8	177	21	AAG48381	Human secreted pro	613	210	22	ABW66738	Novel human diagno
541	5	23.8	177	21	AAG48381	Arabidopsis thalia	614	210	22	AAU25652	G protein-coupled
542	5	23.8	178	22	AAW2073	Dendritic cell (DC	615	211	22	ABW61507	Drosophila melanog
543	5	23.8	179	19	AAW40068	Human monoclonal a	616	212	17	AAW15932	Antibody 7G12 heav
544	5	23.8	179	21	AAG38559	Arabidopsis thalia	617	213	16	AAW4201	Monoclonal antibod
545	5	23.8	181	21	AAG00945	Human monoclonal a	618	213	21	AAW14667	Arabidopsis thalia
546	5	23.8	181	21	AAG11003	Arabidopsis thalia	619	214	20	AAW44176	MAB Fab13B5 heav
547	5	23.8	181	21	AAG47970	Arabidopsis thalia	620	215	16	AAW44176	Monoclonal antibod
548	5	23.8	181	21	AAG47972	Arabidopsis thalia	621	215	16	AAW41747	Human polypeptide
549	5	23.8	181	22	ABG11441	Novel human diagno	622	216	22	AAU03676	Group B streptococ
550	5	23.8	181	22	ABG11441	Human p22 protein	623	217	15	AAW53805	FAB heavy chain fo
551	5	23.8	182	22	ABG19219	Novel human diagno	624	217	16	AAW75460	Mouse antibody H3-
552	5	23.8	183	22	ABW27546	Human peptide #197	625	217	21	AAW18006	Arabidopsis thalia
553	5	23.8	183	22	ABW27546	Peptide #211 encod	626	219	19	AAW53597	Amino acid sequenc
554	5	23.8	183	22	ABW18194	Protein #193 encod	627	220	15	AAW74782	Heavy chain of 59.
555	5	23.8	183	22	AAW53526	Human brain expres	628	221	21	AAW63215	Gene 39 human secr
556	5	23.8	183	22	AAW65905	Human bone marrow	629	221	21	AAW63215	Human secreted pro
557	5	23.8	183	22	AAW13772	Human immune/haema	630	222	15	AAW59415	6D9 antibody heav
558	5	23.8	183	22	AAW26170	Peptide #206 encod	631	222	17	AAW15934	Antibody 3G2 heav
559	5	23.8	183	22	AAW01517	Peptide #207 encod	632	222	22	AAW97388	Spinach precursor
560	5	23.8	184	20	AAW37088	Peptide #199 encod	633	222	22	AAW97388	C glutamicum prote
561	5	23.8	185	22	ABW12284	Amino acid sequenc	634	222	22	AAW92413	Spinach lumazine s
562	5	23.8	185	22	ABW00069	Human secreted pro	635	223	16	AAW75456	Mouse antibody PB3
563	5	23.8	186	21	AAG38538	Human protein SEQ	636	223	16	AAW75456	Mouse antibody F4-
564	5	23.8	186	21	AAG38701	Arabidopsis thalia	637	225	21	AAW22336	Arabidopsis thalia
565	5	23.8	186	21	AAG47774	Arabidopsis thalia	638	225	21	AAW94449	Arabidopsis thalia
566	5	23.8	186	21	AAG48380	Arabidopsis thalia	639	225	22	AAW65272	Human cell surface
567	5	23.8	186	22	ABW64486	Drosophila melanog	640	226	22	ABW67771	Human cell surface
568	5	23.8	187	18	AAV11193	S. pneumoniae phos	641	226	22	AAW98706	Drosophila melanog
569	5	23.8	187	20	AAW59947	Human endometrium	642	227	18	AAW55386	Chandra, a helper
570	5	23.8	187	22	AAW72771	Saccharomyces cere	643	227	18	AAW55386	H. pylori ORF 11ge
571	5	23.8	188	21	AAW95943	Candida albicans p	644	228	21	ABW19236	Novel human diagno
572	5	23.8	189	22	AAW95239	Human reproductive	645	228	21	ABW19236	Human secreted pro
573	5	23.8	189	22	AAW43564	Human polypeptide	646	229	22	AAW18542	Human cytoskeletal
574	5	23.8	189	22	AAW43639	Human polypeptide	647	229	22	AAW38700	Arabidopsis thalia
575	5	23.8	189	22	AAW19951	Novel human calciu	648	230	18	AAW27088	Mouse monoclonal a
576	5	23.8	189	22	AAW74885	Human secreted pro	649	230	21	AAW32942	Pinus radiata tran
577	5	23.8	192	19	AAW91681	Human secreted pro	650	230	22	ABW11220	Novel human diagno
578	5	23.8	192	21	AAW91681	Sequence of new an	651	231	18	AAW27090	Mouse monoclonal a
579	5	23.8	193	7	AAW61520	Novel human diagno	652	231	22	AAW22134	Arabidopsis thalia
580	5	23.8	193	22	ABW08106	Novel human diagno	653	232	22	AAW48673	Propionibacterium
581	5	23.8	193	22	ABW20816	Novel human diagno	654	232	21	AAW69007	Heavy chain of an
582	5	23.8	194	22	ABW53068	Escherichia coli P	655	233	21	AAW69007	Apo-B RNA editing
583	5	23.8	194	21	AAG09995	Arabidopsis thalia	656	236	15	AAW58705	Arabidopsis thalia
584	5	23.8	195	20	AAW26185	Arabidopsis thalia	657	236	22	AAW22335	Human G-protein co
585	5	23.8	196	15	AAW43986	Corn 2-isopropylma	658	237	7	AAW60307	Sequence of new an
586	5	23.8	196	15	AAW43986	p15e Protein. Hum	659	238	14	AAW70306	Human interleukin-
587	5	23.8	196	15	AAW55343	Sequence of envelo	660	238	14	AAW42447	S. glaucosces acb
588	5	23.8	199	22	ABW61521	Sequence of new an	661	238	19	AAW23831	Novel human diagno
589	5	23.8	199	22	ABW28364	Peptide #1015 enco	662	238	22	ABW15868	Human protein sequ
590	5	23.8	199	22	ABW33543	Peptide #1049 enco	663	239	20	AAW43287	HFSH beta subunit-
591	5	23.8	199	22	AAW18999	Protein #998 enco	664	239	20	AAW43287	HFSH beta subunit-
592	5	23.8	199	22	AAW54319	Human brain expres	665	239	20	AAW96745	High affinity immu
593	5	23.8	199	22	AAW66717	Human bone marrow	666	239	21	AAW50174	Human high affinit
594	5	23.8	199	22	AAW14582	Peptide #1016 enco	667	239	22	ABW15553	Novel human diagno
595	5	23.8	199	22	AAW27003	Peptide #1040 enco	668	239	22	AAW74584	Human HA1ERBS Isof
								240	20	AAW43288	HTSH beta subunit-

815	5	23.8	293	21	AAG47235	Arabidopsis thalia	888	23.8	327	21	AAG10038	Arabidopsis thalia
816	5	23.8	293	21	AAG58016	Arabidopsis thalia	889	23.8	327	22	AAB71709	Carp CTR1 protein.
817	5	23.8	294	21	AAG20915	Arabidopsis thalia	890	23.8	327	22	AA877795	Mouse T2R01 amino
818	5	23.8	294	21	AAG37991	Arabidopsis thalia	891	23.8	328	20	AA939999	Human carbonic anhydrase
819	5	23.8	295	7	AAP60326	Interleukin-1 gene	892	23.8	328	20	AA471737	Human PRO237 protein
820	5	23.8	295	7	AAP60680	Prepro-interleukin-1	893	23.8	328	21	AA444293	Human PRO237 protein
821	5	23.8	295	12	AA614855	Mature interleukin-1	894	23.8	328	21	AA24041	Human PRO237 protein
822	5	23.8	295	21	AAG20914	Arabidopsis thalia	895	23.8	328	21	AA19525	Antitumour protein
823	5	23.8	295	21	AAG37990	Arabidopsis thalia	896	23.8	328	21	AA56263	Maedi-Visna virus
824	5	23.8	295	22	AB852467	Arabidopsis thalia	897	23.8	328	22	AAU12334	Human PRO237 polypeptide
825	5	23.8	296	22	AAU37340	Escherichia coli protein	898	23.8	330	22	AB55472	Murine PCB212 protein
826	5	23.8	296	22	AAU22328	Staphylococcus aureus	899	23.8	330	22	AA87811	Mouse T2R17 amino
827	5	23.8	296	22	AAU25335	Human protein sequence	900	23.8	331	21	AA95781	SH7 single chain antibody
828	5	23.8	297	17	AAG82937	S. epidermidis opes	901	23.8	331	22	AA92926	Novel human enzyme
829	5	23.8	297	17	AAW00283	Mutant farnesylidp	902	23.8	332	21	AA80583	Arabidopsis thalia
830	5	23.8	298	12	AAW00286	Native farnesylidp	903	23.8	332	21	AA806448	Arabidopsis thalia
831	5	23.8	298	21	AAG84912	Shrimp white spot	904	23.8	332	21	AA806484	Arabidopsis thalia
832	5	23.8	299	21	AAG05420	Arabidopsis thalia	905	23.8	332	21	AA807390	Arabidopsis thalia
833	5	23.8	299	21	AAG47773	Arabidopsis thalia	906	23.8	332	21	AA807390	Arabidopsis thalia
834	5	23.8	299	20	AA91352	Human secreted protein	907	23.8	332	21	AA824178	Arabidopsis thalia
835	5	23.8	300	20	AA941759	Human PRO1014 protein	908	23.8	332	21	AA833349	Ze mays protein f
836	5	23.8	300	21	AA844315	Human PRO1014 (UNQ)	909	23.8	332	21	AA834599	Arabidopsis thalia
837	5	23.8	300	22	AAU22328	Novel human enzyme	910	23.8	332	21	AA837213	Arabidopsis thalia
838	5	23.8	300	22	AAU29089	Human PRO polypeptide	911	23.8	332	21	AA837317	Arabidopsis thalia
839	5	23.8	301	22	AA819930	Human oxidoreductase	912	23.8	332	21	AA837989	Arabidopsis thalia
840	5	23.8	301	21	AA818895	A maize chitinase	913	23.8	332	21	AA841511	Arabidopsis thalia
841	5	23.8	302	21	AAG14666	Arabidopsis thalia	914	23.8	332	21	AA844697	Ze mays protein f
842	5	23.8	302	21	AA854312	Human pancreatic c	915	23.8	332	21	AA847761	Arabidopsis thalia
843	5	23.8	302	22	AAU49983	Arabidopsis thalia	916	23.8	332	21	AA847770	Arabidopsis thalia
844	5	23.8	302	22	AAU16378	Human novel secret	917	23.8	332	21	AA847770	Arabidopsis thalia
845	5	23.8	303	22	AAG90217	C glutamicum prote	918	23.8	332	21	AA850873	Arabidopsis thalia
846	5	23.8	303	18	AAW24590	H. pylori derived	919	23.8	332	21	AA850944	Arabidopsis thalia
847	5	23.8	303	18	AAW20132	H. pylori derived	920	23.8	333	21	AA843046	Human OREF ORF2810
848	5	23.8	303	21	AAG10321	Arabidopsis thalia	921	23.8	333	22	AA842076	Human polypeptide
849	5	23.8	303	22	AAG90813	C glutamicum prote	922	23.8	333	22	AA82048	Human G protein co
850	5	23.8	304	19	AAW40070	Mouse H1L1 amino	923	23.8	335	22	AB868367	Drosophila melanog
851	5	23.8	304	22	AAU07621	Mouse 2P channel p	924	23.8	335	22	AA936088	Klebsiella pneumonia
852	5	23.8	304	22	AAE00947	Human monoclonal a	925	23.8	335	22	AA933358	Human polypeptide
853	5	23.8	305	22	AAU37698	Streptococcus pneu	926	23.8	337	20	AA9700281	Human secreted pro
854	5	23.8	305	22	AAU37960	Streptococcus pneu	927	23.8	337	22	AA837208	Arabidopsis thalia
855	5	23.8	307	22	AAU91421	Human secreted pro	928	23.8	337	22	AAU52875	Propionibacterium
856	5	23.8	307	22	AAU33659	Pseudomonas aerugi	929	23.8	337	22	AA867776	Amino acid sequenc
857	5	23.8	308	21	AAU37210	Arabidopsis thalia	930	23.8	339	20	AA9367799	Amino acid sequenc
858	5	23.8	308	21	AAU05419	Arabidopsis thalia	931	23.8	339	20	AA93406	Novel human diagno
859	5	23.8	309	21	AAG57419	Arabidopsis thalia	932	23.8	339	22	AA826859	Thermopsis. Sulph
860	5	23.8	309	21	AAG47772	Arabidopsis thalia	933	23.8	340	11	AA806830	Sequence encoded b
861	5	23.8	310	22	ABG05624	Salmonella typhi c	934	23.8	341	9	AA853200	Arabidopsis thalia
862	5	23.8	311	20	AAW90348	Human FAST-1 prote	935	23.8	341	22	AA804645	Human polypeptide
863	5	23.8	311	22	AA837209	Arabidopsis thalia	936	23.8	341	22	AA843550	Human polypeptide
864	5	23.8	311	22	AB871370	Drosophila melanog	937	23.8	341	22	AA843628	Human polypeptide
865	5	23.8	313	22	AA814138	Bordetella pertuss	938	23.8	341	22	AAU19933	Novel human calciu
866	5	23.8	313	22	AAU91459	Human secreted pro	939	23.8	342	21	AA861158	Arabidopsis thalia
867	5	23.8	313	21	AAU91459	Human secreted pro	940	23.8	342	22	AAU01250	B. subtilis acetoh
868	5	23.8	314	21	AAU95770	Human olfactory re	941	23.8	342	22	AA827240	Human EXMAD-18 SEQ
869	5	23.8	316	20	AAV19508	Human prostate can	942	23.8	345	22	AA827464	Arabidopsis thalia
870	5	23.8	316	22	AA887803	Amino acid sequenc	943	23.8	345	21	AA827464	Arabidopsis thalia
871	5	23.8	317	22	AB869511	Mouse T2R09 amino	944	23.8	345	21	AA810571	Arabidopsis thalia
872	5	23.8	318	22	AB869511	Drosophila melanog	945	23.8	345	21	AA810571	Arabidopsis thalia
873	5	23.8	319	17	AA898413	Arabidopsis thalia	946	23.8	345	21	AA847748	Arabidopsis thalia
874	5	23.8	319	22	AA898413	Mouse myeloma MOPC	947	23.8	345	22	AA849690	Human colon cancer
875	5	23.8	321	22	AA871491	Human olfactory re	948	23.8	347	14	AA834224	Arabidopsis thalia
876	5	23.8	321	22	AA871491	Mouse myeloma MOPC	949	23.8	347	14	AA834224	Arabidopsis thalia
877	5	23.8	322	21	AA878311	Human PTB1B mutat	950	23.8	348	20	AA937541	Amino acid sequenc
878	5	23.8	322	21	AA878311	Human PTB1B mutat	951	23.8	348	20	AA937541	Human polypeptide
879	5	23.8	322	21	AA878311	Arabidopsis thalia	952	23.8	348	22	AA871490	Protein which is s
880	5	23.8	322	21	AA878311	Arabidopsis thalia	953	23.8	348	22	AA871490	Human olfactory re
881	5	23.8	323	17	AA898412	Arabidopsis thalia	954	23.8	348	22	AA872907	Human olfactory re
882	5	23.8	323	19	AA898412	Mouse immunoglobul	955	23.8	354	22	AA809788	P. papuli xylogluc
883	5	23.8	323	19	AA898412	Mutant human PTP1B	956	23.8	354	21	AA848893	Arabidopsis thalia
884	5	23.8	323	19	AA898412	Mutant human PTP1B	957	23.8	354	21	AA848893	Eucalyptus grandis
885	5	23.8	324	21	AAU24715	Human olfactory re	958	23.8	359	22	AA832709	Canine retrovirus
886	5	23.8	324	21	AAU10039	Arabidopsis thalia	959	23.8	359	22	AA849469	Arabidopsis thalia
887	5	23.8	324	21	AAU10039	Newborn mouse immu	960	23.8	360	21	AA827740	Arabidopsis thalia
888	5	23.8	324	22	AAU10039	Propionibacterium	961	23.8	360	21	AA805382	Arabidopsis thalia
889	5	23.8	324	22	AAU10039	Murine OR-like pol	962	23.8	360	21	AA805382	Arabidopsis thalia
890	5	23.8	324	22	AAU10039	Murine OR-like pol	963	23.8	360	21	AA805382	Arabidopsis thalia

961 Arabidopsis thalia
962 Arabidopsis thalia
963 Zea mays protein f
964 Arabidopsis thalia
965 Arabidopsis thalia
966 Arabidopsis thalia
967 Arabidopsis thalia
968 Arabidopsis thalia
969 Arabidopsis thalia
970 Zea mays protein f
971 Arabidopsis thalia
972 Arabidopsis thalia
973 Arabidopsis thalia
974 Arabidopsis thalia
975 Arabidopsis thalia
976 Arabidopsis thalia
977 Zea mays protein f
978 Arabidopsis thalia
979 Arabidopsis thalia
980 Arabidopsis thalia
981 Arabidopsis thalia
982 Arabidopsis thalia
983 Arabidopsis thalia
984 Arabidopsis thalia
985 Arabidopsis thalia
986 Arabidopsis thalia
987 Arabidopsis thalia
988 Arabidopsis thalia
989 Arabidopsis thalia
990 Arabidopsis thalia
991 Arabidopsis thalia
992 Arabidopsis thalia
993 Arabidopsis thalia
994 Arabidopsis thalia
995 Arabidopsis thalia
996 Arabidopsis thalia
997 Arabidopsis thalia
998 Arabidopsis thalia
999 Arabidopsis thalia
1000 Arabidopsis thalia

ALIGNMENTS

RESULT 1

ID ABB70857 standard; Protein; 929 AA.

AC ABB70857;
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster polypeptide SEQ ID NO 39363.
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.
DR N-PSDB; ABL14960.

XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

XX Disclosure; SEQ ID NO 39363; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA sequences (AB16176-ABL30511) and the encoded proteins (AB16176-ABL30511).

XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 929 AA;

Query Match 33.3%; Score 7; DB 22; Length 929;

Best Local Similarity 100.0%; Pred. No. 58;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SFTQALG 20

DB 103 SFTQALG 109

RESULT 2

ID AAY40080 standard; Peptide; 27 AA.

AC AAY40080;

DT 18-NOV-1999 (first entry)

DE Peptide sequence derived from a human secreted protein.

XX Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
KW neurodegenerative disorder; developmental abnormality; blood disorder;
KW immune system disease; autoimmune disease; leukemia; inflammation;
KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
KW connective tissue disorder; transplant rejection; sepsis; acne;
KW psoriasis; cardiovascular disorder; reproductive disorder;
KW food additive; food preservative; storage capability.

XX Homo sapiens.

XX WO9943693-A1.

XX 02-SEP-1999.

XX 24-FEB-1999; 99WO-US03939.

XX 26-FEB-1998; 98US-0076051.

XX 26-FEB-1998; 98US-0076052.

XX 26-FEB-1998; 98US-0076053.

XX 26-FEB-1998; 98US-0076054.

XX 26-FEB-1998; 98US-0076057.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;

PI Duan RD;

XX WPI; 1999-550857/46.

XX New human genes and the secreted polypeptides they encode, useful for

PT diagnosis and treatment of e.g. cancers, neurological disorders, immune
 PT diseases, inflammation or blood disorders
 XX Disclosure; Page 53; 246pp; English.
 PS
 CC RAY40001-92 are derived from human secreted proteins. The
 CC polynucleotides and their corresponding secreted polypeptides are useful
 CC for preventing, treating or ameliorating medical conditions, e.g. by
 CC protein or gene therapy. Pathological conditions can also be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the polynucleotide. Specific
 CC uses include developing products for the diagnosis or treatment of
 CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
 CC and fetal deficiencies, blood disorders, sepsis, diseases of the immune
 CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and
 CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
 CC infections, AIDS, connective tissue disorders, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC and reproductive disorders. The polypeptides or polynucleotides can
 CC also be used as food additives or preservatives, such as to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components.
 XX
 SQ Sequence 27 AA;
 Query Match 28.6%; Score 6; DB 20; Length 27;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 15 FTQALG 20
 Db 21 FTQALG 26
 RESULT 3
 AAM23675
 ID AAM23675 standard; Protein; 48 AA.
 AC AAM23675;
 XX
 DT 12-OCT-2001 (first entry)
 XX Human EST encoded protein SEQ ID NO: 1200.
 DE
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder;
 KW biodiversity; gene therapy; nutrition.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 25-JAN-2001; 2001WO-US02687.
 XX
 XX 25-JAN-2000; 2000US-0491404.
 PR 17-JUL-2000; 2000US-0617746.
 PR 03-AUG-2000; 2000US-0631451.
 PR 15-SEP-2000; 2000US-0663870.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 XX WPI: 2001-476164/51.
 DR N-PSDB; AAH98334.
 DR
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use -

XX Claim 20; Page 868; 1275pp; English.
 PS
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a
 CC protein of the invention.
 XX
 SQ Sequence 48 AA;
 Query Match 28.6%; Score 6; DB 22; Length 48;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SGLSLT 6
 Db 36 SGLSLT 41
 RESULT 4
 AAR94979
 ID AAR94979 standard; Protein; 51 AA.
 AC AAR94979;
 XX
 DT 05-JAN-1997 (first entry)
 XX Calcium channel isoform CaCh3 IVS3-IVS4 region (ROB3).
 DE
 XX Stretch-activated cation channel; SA-Cat; calcium channel; CaCh3;
 KW bone; osteoblast; antisense; osteosclerosis; hypertension.
 KW
 XX Rattus sp.
 OS
 XX
 PH Key
 FT Domain
 FT 1-19
 FT /label= IVS3
 FT /note= "transmembrane domain 3"
 FT 33-51
 FT /label= IVS4
 FT /note= "transmembrane domain 4"
 FT
 XX WO9613269-A1.
 XX
 XX 09-MAY-1996.
 PD
 XX 25-OCT-1995; 95WO-US13686.
 PF
 XX 28-OCT-1994; 94US-0330433.
 PR
 XX (DART-) DARTMOUTH COLLEGE.
 PA (JEWI-) JEWISH HOSPITAL ST LOUIS.
 XX
 XX Barry ELR, Duncan RL, Friedman PA, Hruska KA;
 PI
 XX WPI: 1996-239267/24.
 DR N-PSDB; AAT18887.
 DR
 XX Stretch activated cation channel gene antisense oligonucleotide -
 PT used in the treatment of hypertension and osteosclerosis
 XX
 XX Example 1; Fig 4; 60pp; English.
 PS
 XX Amino acid sequences (AAR94977-79) of UMR-106 clones ROB1, ROB2 and
 CC ROB3 (see also AAT44381 and AAT18886-87) were deduced for the region
 CC including transmembrane domains 3 and 4. ROB1, ROB2 and ROB3 correspond
 CC to isoforms CaCh1, CaCh2 and CaCh3, respectively, of domain 4 of the
 CC alpha 1 subunit of a stretch-activated cation (SA-Cat) channel.
 CC Comparison of the sequences with corresponding sequences of rat.

CC brain L-type calcium channels show that CaCh from rat osteosarcoma
 CC UMR-106 cells lack a portion of the IVS3-IV4 linker domain as a
 CC result of alternative splicing.

XX SQ Sequence 51 AA;

Query Match 28.6%; Score 6; DB 17; Length 51;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
 |||||
 DB 34 TFFRLF 39

RESULT 5

AAR94978
 ID AAR94978 standard; Protein; 51 AA.

XX AAR94978;
 AC
 XX

DT 05-JAN-1997 (first entry)

XX Calcium channel isoform CaCh2 IVS3-IVS4 region (ROB2).

DE Stretch-activated cation channel; SA-Cat; calcium channel; CaCh2;
 KW bone; osteoblast; antisense; osteosclerosis; hypertension.
 XX

OS Rattus sp.

XX Key Location/Qualifiers
 FT Domain 1..19

FT /label= IVS3
 FT /note= "transmembrane domain 3"
 FT 33..51

FT /label= IVS4
 FT /note= "transmembrane domain 4"

XX WO9613269-A1.

PW
 XX

PD 09-MAY-1996.

XX 25-OCT-1995; 95WO-US13686.

XX 28-OCT-1994; 94US-0330433.

XX (DART-) DARTMOUTH COLLEGE.
 XX (JEWI-) JEWISH HOSPITAL ST LOUIS.

PI Barry ELR, Duncan RL, Friedman PA, Hruska KA;

XX WPI; 1996-239267/24.
 DR N-PSDB; AAT18886.

XX Stretch activated cation channel gene antisense oligonucleotide
 used in the treatment of hypertension and osteosclerosis

PS Example 1; Fig 4; 60pp; English.

XX Amino acid sequences (AAR94977-79) of UMR-106 clones ROB1, ROB2 and
 CC ROB3 (see also AAT44381 and AAT18886-87) were deduced for the region
 CC including transmembrane domains 3 and 4. ROB1, ROB2 and ROB3 correspond
 CC to isoforms CaCh1, CaCh2 and CaCh3, respectively, of domain 4 of the
 CC alpha 1 subunit of a stretch-activated cation (SA-Cat) channel.
 CC Comparison of the sequences with corresponding sequences of rat
 CC brain L-type calcium channels show that CaCh from rat osteosarcoma
 CC UMR-106 cells lack a portion of the IVS3-IV4 linker domain as a
 CC result of alternative splicing. ROB2 is a close isoform of the
 CC SA-Cat channel.

XX SQ Sequence 51 AA;

Query Match 28.6%; Score 6; DB 17; Length 51;

Best Local Similarity 100.0%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
 |||||
 DB 34 TFFRLF 39

RESULT 6

AAR97851
 ID AAR97851 standard; Protein; 62 AA.

XX AAR97851;
 AC
 XX

DT 05-JAN-1997 (first entry)

XX Rat brain calcium channel IVS3-IVS4 region (RatBr2).

DE Stretch-activated cation channel; SA-Cat; calcium channel; CaCh;
 KW bone; osteoblast; antisense; osteosclerosis; hypertension.
 XX

OS Rattus sp.

XX Key Location/Qualifiers
 FT Domain 1..19

FT /label= IVS3
 FT /note= "transmembrane domain 3"
 FT 44..62

FT /label= IVS4
 FT /note= "transmembrane domain 4"

XX WO9613269-A1.

PW
 XX

PD 09-MAY-1996.

XX 25-OCT-1995; 95WO-US13686.

XX 28-OCT-1994; 94US-0330433.

XX (DART-) DARTMOUTH COLLEGE.
 XX (JEWI-) JEWISH HOSPITAL ST LOUIS.

PI Barry ELR, Duncan RL, Friedman PA, Hruska KA;

XX WPI; 1996-239267/24.

XX Stretch activated cation channel gene antisense oligonucleotide
 used in the treatment of hypertension and osteosclerosis

PS Example 1; Fig 4; 60pp; English.

XX Amino acid sequences (AAR97850-52) for the transmembrane domain
 CC 3 to 4 region of rat brain calcium channel proteins RabSkell,
 CC RatBr2 and RatBr3, respectively, were compared to corresponding
 CC regions of calcium channel CaCh1 (ROB1), CaCh2 (ROB2) and CaCh3
 CC (ROB3) isoforms (see also AAR94977-79) deduced from rat osteosarcoma
 CC UMR-106 cDNA clones (AAT44381 and AAT18886-87). The comparison revealed
 CC the use of an alternative splice acceptor site or an exon skipping
 CC event in the IVS3-IVS4 linker, producing shorter transcripts in ROB1-3.

XX SQ Sequence 62 AA;

Query Match 28.6%; Score 6; DB 17; Length 62;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
 |||||
 DB 45 TFFRLF 50

RESULT 7

AAR97852

ID AAR97852 standard; Protein; 66 AA.
 AC AAR97852;
 XX
 XX
 DT 05-JAN-1997 (first entry)
 XX
 XX Rat brain calcium channel IVS3-IVS4 region (RatBr3).
 DE
 XX Stretch-activated cation channel; SA-Cat; calcium channel; CaCh;
 KW bone; osteoblast; antisense; osteosclerosis; hypertension.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..19
 FT /label= IVS3
 FT /note= "transmembrane domain 3"
 FT 48..66
 FT /label= IVS4
 FT /note= "transmembrane domain 4"
 FT
 XX WO9613269-A1.
 PN
 XX
 PD 09-MAY-1996.
 XX
 XX 25-OCT-1995; 95WO-US13686.
 PF
 XX 28-OCT-1994; 94US-0330433.
 PR
 XX (DART-) DARTMOUTH COLLEGE.
 PA (JEWI-) JEWISH HOSPITAL ST LOUIS.
 XX
 XX Barry ELR, Duncan RL, Friedman PA, Hruska KA;
 PI WPI; 1996-239267/24.
 DR
 XX Stretch activated cation channel gene antisense oligonucleotide
 PT used in the treatment of hypertension and osteosclerosis
 XX
 PS Example 1; Fig 4; 60pp; English.
 XX
 XX Amino acid sequences (AAR97850-52) for the transmembrane domain
 CC 3 to 4 region of rat brain calcium channel proteins RabSkell,
 CC RatBr2 and RatBr3, respectively, were compared to corresponding
 CC regions of calcium channel CaCh1 (ROB1), CaCh2 (ROB2) and CaCh3
 CC (ROB3) isoforms (see also AAR94977-79) deduced from rat osteosarcoma
 CC UMR-106 cDNA clones (AAT44381 and AAT18886-87). The comparison revealed
 CC the use of an alternative splice acceptor site or an exon skipping
 CC event in the IVS3-IVS4 linker, producing shorter transcripts in ROB1-3.
 XX
 SQ Sequence 66 AA;
 Query Match 28.6%; Score 6; DB 17; Length 66;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 TFFRLF 11
 Db 49 TFFRLF 54
 RESULT 8
 ABG04703
 ID ABG04703 standard; Protein; 66 AA.
 XX
 AC ABG04703;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #4694.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US08631.
 PF
 XX 31-MAR-2000; 2000US-0540217.
 PR
 XX 23-AUG-2000; 2000US-0649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 XX
 XX WPI; 2001-639362/73.
 DR N-PSDB; AAS68890.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 PS Claim 20; SEQ ID NO 35062; 103pp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 66 AA;
 Query Match 28.6%; Score 6; DB 22; Length 66;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GSLSTF 7
 Db 28 GSLSTF 33
 RESULT 9
 AAM89377
 ID AAM89377 standard; Protein; 78 AA.
 XX
 AC AAM89377;
 XX
 DT 07-NOV-2001 (first entry)
 XX
 XX Human immune/haematopoietic antigen SEQ ID NO:16970.
 DE
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis.
 XX
 OS Homo sapiens.

[illegible]

PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-483426/52.
 DR N-PSDB; AAK62158.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Claim 11; SEQ ID NO 16970; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 78 AA;
 Query Match 28.6%; Score 6; DB 22; Length 78;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 GSLSTF 7
 Db 22 GSLSTF 27
 |||||
 RESULT 10
 AAW59796
 ID AAW59796 standard; Protein: 96 AA.
 XX
 AC AAW59796;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Nucleotide sequence of complete EST3 gene product 2.
 XX
 KW EST3 gene; telomerase inhibitor; cancer; mammal; tumour growth;
 KW stem cell; chemotherapeutic agent.
 XX
 OS Saccharomyces cerevisiae.
 XX
 PN W09823759-A2.
 XX
 PD 04-JUN-1998.
 XX
 PF 26-NOV-1997; 97WO-US21272.
 XX
 PR 26-NOV-1996; 96US-0756693.
 XX
 PA (BAYU) BAYLOR COLLEGE MEDICINE.
 XX
 PI Lundblad V;
 XX
 XX WPI: 1998-332927/29.
 DR N-PSDB; AAV41581.
 XX
 PT Telomerase-associated proteins useful to identify telomerase
 PT inhibitors - useful e.g. for cancer treatment in mammals, and
 PT screening methods to isolate additional telomerase-associated genes

PT or protein homologues
 XX Disclosure: Fig 7; 7lpp; English.
 XX
 CC This is the nucleotide sequence of the EST3 gene product 2 (including
 CC flanking regions and the +1 ribosomal frame shift), used in the method
 CC of the invention to identify telomerase inhibitors which are useful in
 CC the treatment of cancer. The proteins are useful to isolate
 CC telomerase-inhibiting compounds. Such inhibitors are useful in cancer
 CC treatment in mammals, since reactivation of telomerase (normally present
 CC only in germ line cells) is thought to be necessary for sustained tumour
 CC growth, and only tumour and stem cells would be targeted by such agents,
 CC producing limited side effects compared to chemotherapeutic agents.
 XX
 SQ Sequence 96 AA;
 Query Match 28.6%; Score 6; DB 19; Length 96;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 2 GSLSTF 7
 Db 84 GSLSTF 89
 |||||
 RESULT 11
 AAY40074
 ID AAY40074 standard; Peptide: 122 AA.
 XX
 AC AAY40074;
 XX
 DT 18-NOV-1999 (first entry)
 XX
 DE Peptide sequence derived from a human secreted protein.
 XX
 KW Secreted protein; gene therapy; cancer; tumor; fetal deficiency;
 KW neurodegenerative disorder; developmental abnormality; blood disorder;
 KW immune system disease; autoimmune disease; leukemia; inflammation;
 KW allergy; Alzheimer's disease; cognitive disorder; schizophrenia;
 KW obesity; osteoporosis; arthritis; infection; AIDS; diabetes; asthma;
 KW connective tissue disorder; transplant rejection; sepsis; acne;
 KW psoriasis; cardiovascular disorder; reproductive disorder;
 KW food additive; food preservative; storage capability.
 XX
 OS Homo sapiens.
 XX
 PN W09943693-A1.
 XX
 PD 02-SEP-1999.
 XX
 PF 24-FEB-1999; 99WO-US03939.
 XX
 PR 26-FEB-1998; 98US-0076051.
 XX
 PR 26-FEB-1998; 98US-0076052.
 XX
 PR 26-FEB-1998; 98US-0076053.
 XX
 PR 26-FEB-1998; 98US-0076054.
 XX
 PR 26-FEB-1998; 98US-0076057.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Olsen HS, Florence K, Brewer LA, Ebner R, Ruben SM, Rosen CA;
 PI Duan RD;
 XX
 XX WPI: 1999-550857/46.
 XX
 PT New human genes and the secreted polypeptides they encode, useful for
 PT diagnosis and treatment of e.g. cancers, neurological disorders, immune
 PT diseases, inflammation or blood disorders
 XX
 PS Disclosure: Page 53; 246pp; English.
 XX
 CC AAY40001-92 are derived from human secreted proteins. The
 CC polynucleotides and their corresponding secreted polypeptides are useful

CC for preventing, treating or ameliorating medical conditions, e.g. by
 CC protein or gene therapy. Pathological conditions can also be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC determining the presence of mutations in the polynucleotide. Specific
 CC uses include developing products for the diagnosis or treatment of
 CC cancer, tumors, neurodegenerative disorders, developmental abnormalities
 CC and fetal deficiencies, blood disorders, sepsis, diseases of the immune
 CC system, autoimmune diseases, inflammation, allergies, Alzheimer's and
 CC cognitive disorders, schizophrenia, obesity, osteoporosis, arthritis,
 CC infections, AIDS, connective tissue disorders, transplant rejection,
 CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
 CC and reproductive disorders. The polypeptides or polynucleotides can
 CC also be used as food additives or preservatives, such as to increase
 CC or decrease storage capabilities, fat content, lipid, protein,
 CC carbohydrate, vitamins, minerals, cofactors or other nutritional
 CC components.
 XX
 SQ Sequence 122 AA;

Query Match 28.6%; Score 6; DB 20; Length 122;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 FTQALG 20
 Db 97 FTQALG 102
 |||||

RESULT 12

AAU45360

ID AAU45360 standard; Protein; 141 AA.
 XX
 AC AAU45360;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #6256.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

XX L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59525.

XX Propionibacterium acnes polypeptides and nucleic acids useful for

XX vaccinating against and diagnosing infections, especially useful for

XX treating acne vulgaris.

XX Example 1; SEQ ID No 6555; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX polypeptides. The proteins and their associated DNA sequences are used in

XX the treatment, prevention and diagnosis of medical conditions caused by

XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,

CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 141 AA;

Query Match 28.6%; Score 6; DB 22; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SFTQAL 19
 Db 47 SFTQAL 52
 |||||

RESULT 13

AAE11946

ID AAE11946 standard; Protein; 147 AA.

XX
 AC AAE11946;

XX 18-DEC-2001 (first entry)

XX Human CCAAT/enhancer binding protein (C/EBP) beta-3 or p20 isoform.

XX Human; CCAAT/enhancer binding protein; C/EBPbeta; transcription factor;
 KW interleukin; IL; p20; inflammation; adult respiratory distress syndrome;
 KW allergic rhinitis; arthritis; bronchitis; bronchopulmonary dysplasia;
 KW cystic fibrosis; extensive allergic alveolitis; anti-inflammatory;
 KW idiopathic pulmonary fibrosis; interstitial lung disease; anti-allergic;
 KW inflammatory bowel disease; respiratory viral infection; anti-arthritis;
 KW anti-asthma; intestinal; antiviral.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 66..82

XX /label= NLS_A

XX /note= "Nuclear localisation sequence A"

XX Modified-site 68

XX /note= "Phosphorylation site"

XX Peptide 79..95

XX /label= NLS_B

XX /note= "Nuclear localisation sequence B"

XX WO200160320-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05578.

XX 18-FEB-2000; 2000US-0183584.

XX (UVA-) UNIV VANDERBILT.

XX Bringham KL, Stecenko AA, Sealy L;

XX WPI; 2001-581897/65.

XX N-PSDB; AAD19382.

PT Treating inflammation, particularly of the lung, by increasing activity
 PT of p20, the beta3-isoform of CCAAT/enhancer binding protein -
 Example 5; Fig 5; 200pp; English.
 PS
 XX
 CC The present sequence is a human CCAAT/Enhancer Binding Protein (C/EBP)
 CC beta3 (referred as p20) isoform. The C/EBPbeta is a transcription factor
 CC which is identified as being critical for maximal interleukin (IL)-6 and
 CC IL-8 expression. The isoforms of C/EBPbeta are C/EBPbeta-1, C/EBPbeta-2
 CC and C/EBPbeta-3. The p20 isoform of C/EBPbeta is useful for treating
 CC inflammation, adult respiratory distress syndrome, allergic rhinitis,
 CC arthritis, bronchitis, bronchopulmonary dysplasia, cystic fibrosis,
 CC extensive allergic alveolitis, idiopathic pulmonary fibrosis,
 CC inflammatory bowel disease, interstitial lung disease and respiratory
 CC viral infection.
 XX
 XX
 SQ Sequence 147 AA;

Query Match 28.6%; Score 6; DB 22; Length 147;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGLST 6
 Db 23 SGLST 28
 |||||

RESULT 14
 AAU16099
 ID AAU16099 standard; Proteins; 150 AA.

XX AC AAU16099;
 XX DT 07-NOV-2001 (first entry)
 XX DE Human novel secreted protein, Seq ID 1052.

XX Human; immunosuppressive; antiarthritic; antirheumatic;
 KW cytostatic; cardiant; vasotropic; cerebroprotective; nootropic;
 KW neuroprotective; antibacterial; virucide; fungicide; ophthalmological;
 KW vulnery; secreted protein; rheumatoid arthritis;
 KW hyperproliferative disorder; cardiovascular disorder; cardiac arrest;
 KW cerebrovascular disorder; cerebral ischaemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; infection; ocular disorder;
 KW corneal infection; wound healing; epithelial cell proliferation;
 KW skin ageing; food additive; preservative; antiproliferative.

XX Homo sapiens.
 XX OS
 XX PN WC200155322-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01341.

XX 31-JAN-2000; 2000US-0179065.
 XX 04-FEB-2000; 2000US-0180628.
 XX 24-FEB-2000; 2000US-0184664.
 XX 02-MAR-2000; 2000US-0186350.
 XX 16-MAR-2000; 2000US-0189874.
 XX 17-MAR-2000; 2000US-0190076.
 XX 18-APR-2000; 2000US-0198123.
 XX 19-MAY-2000; 2000US-0205515.
 XX 07-JUN-2000; 2000US-0209467.
 XX 28-JUN-2000; 2000US-0214886.
 XX 30-JUN-2000; 2000US-0215135.
 XX 07-JUL-2000; 2000US-0216647.
 XX 07-JUL-2000; 2000US-0216880.
 XX 11-JUL-2000; 2000US-0217487.
 XX 11-JUL-2000; 2000US-0217496.
 XX 14-JUL-2000; 2000US-0218290.
 XX 26-JUL-2000; 2000US-0220963.
 XX 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225457.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235835.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0245477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250191.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.

DR N-PSDB; AAS26086.

DR New nucleic acid molecules encoding 461 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -

XX Claim 11; SEQ ID No 1052; 960pp; English.

XX The invention relates to isolated nucleic acid molecules and their
 CC encoded secreted proteins. The nucleic acids and proteins are used to
 CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. They
 CC are also used in diagnosing a pathological condition or susceptibility
 CC to a pathological condition. Antibodies to the proteins can also
 CC be used in alleviating symptoms associated with the disorders and in
 CC diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
 CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated
 CC include autoimmune diseases e.g. rheumatoid arthritis,
 CC hyperproliferative disorders e.g. neoplasms of the breast or liver,
 CC cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders
 CC e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g.
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi

CC and ocular disorders e.g. corneal infection, and many other
 CC disorders listed in the specification. The polypeptides can also
 CC be used to aid wound healing and epithelial cell proliferation, to
 CC prevent skin aging due to sunburn, to maintain organs before
 CC transplantation, for supporting cell culture of primary tissues, to
 CC regenerate tissues and in chemotaxis. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
 CC minerals, cofactors and other nutritional components. The present
 CC sequence represents a novel secreted protein of the invention.

Query Match 28.6%; Score 6; DB 22; Length 150;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RSFTQA 18
 |||||
 Db 119 RSFTQA 124

RESULT 15

AAG48562

ID AAG48562 standard; Protein; 152 AA.

XX AC AAG48562;

XX DT 18-OCR-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61338.

XX KW Protein identification; signal transduction pathway; metabolic pathway;

XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR	24-MAY-1999;	99US-0135621.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137503.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140699.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	20-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149802.
PR	23-AUG-1999;	99US-0149930.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	30-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	22-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	28-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.

PR 28-OCT-1999; 99US-01611992.
PR 28-OCT-1999; 99US-01611993.
PR 29-OCT-1999; 99US-0162142.

Query Match 28.6%; Score 6; DB 21; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SCSLST 6
|||||
Db 84 SCSLST 89

Search completed: November 5, 2002, 11:05:26
Job time : 44.6119 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 11:03:27 ; Search time 6.89552 Seconds
(without alignments)
74.387 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 21

Sequence: 1 SGLSTFFRLNRSFTQALCK 21

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	28.6	282	4	US-08-482-918-56
2	6	28.6	282	4	US-09-224-681-56
3	6	28.6	282	4	US-08-336-728A-56
4	6	28.6	446	4	US-09-268-992-51
5	6	28.6	466	4	US-09-268-992-64
6	6	28.6	477	4	US-09-268-992-4
7	6	28.6	495	4	US-09-268-992-2
8	6	28.6	499	2	US-07-952-853-6
9	6	28.6	499	2	US-08-914-848-6
10	6	28.6	635	2	US-09-014-969-11
11	6	28.6	1968	1	US-07-745-206A-7
12	6	28.6	1968	1	US-08-455-543A-45
13	6	28.6	1968	2	US-08-223-305C-45
14	6	28.6	1968	1	US-08-311-363-7
15	6	28.6	2161	1	US-07-745-206A-2
16	6	28.6	2161	1	US-08-455-543A-49
17	6	28.6	2161	1	US-08-455-543A-51
18	6	28.6	2161	2	US-08-223-305C-49
19	6	28.6	2161	2	US-08-311-363-2
20	5	23.8	8	4	US-09-139-802-4
21	5	23.8	15	2	US-08-726-464B-30
22	5	23.8	16	2	US-08-480-190-261
23	5	23.8	16	2	US-08-488-379-261
24	5	23.8	16	5	PCT-US93-07545-261
25	5	23.8	23	1	US-08-383-753-82
26	5	23.8	23	2	US-08-586-772-82
27	5	23.8	23	2	US-08-586-772-82

28	5	23.8	23	2	US-08-959-512-82	Sequence 82, Appl
29	5	23.8	23	4	US-09-512-983-82	Sequence 82, Appl
30	5	23.8	31	4	US-09-177-249-21	Sequence 21, Appl
31	5	23.8	73	2	US-08-245-511-36	Sequence 36, Appl
32	5	23.8	73	2	US-08-600-993A-36	Sequence 36, Appl
33	5	23.8	75	3	US-09-042-012-8	Sequence 8, Appl
34	5	23.8	75	4	US-09-305-086-2	Sequence 2, Appl
35	5	23.8	75	4	US-09-457-324-8	Sequence 8, Appl
36	5	23.8	78	1	US-08-680-726A-60	Sequence 60, Appl
37	5	23.8	78	4	US-09-092-409-60	Sequence 60, Appl
38	5	23.8	97	3	US-08-300-928C-10	Sequence 10, Appl
39	5	23.8	97	3	US-08-430-944D-10	Sequence 10, Appl
40	5	23.8	97	3	US-08-430-014-10	Sequence 10, Appl
41	5	23.8	97	3	US-08-431-184-10	Sequence 10, Appl
42	5	23.8	97	4	US-08-858-207A-431	Sequence 431, App
43	5	23.8	109	1	US-07-662-193-5	Sequence 5, Appl
44	5	23.8	109	1	US-07-807-529A-6	Sequence 6, Appl
45	5	23.8	109	3	US-08-300-928C-8	Sequence 8, Appl
46	5	23.8	109	3	US-08-430-944D-8	Sequence 8, Appl
47	5	23.8	109	3	US-08-430-014-8	Sequence 8, Appl
48	5	23.8	109	3	US-08-431-184-8	Sequence 8, Appl
49	5	23.8	109	5	PCT-US93-02462-6	Sequence 6, Appl
50	5	23.8	111	1	US-07-662-193-4	Sequence 4, Appl
51	5	23.8	111	3	US-08-300-928C-6	Sequence 6, Appl
52	5	23.8	111	3	US-08-430-944D-6	Sequence 6, Appl
53	5	23.8	111	3	US-08-430-014-6	Sequence 6, Appl
54	5	23.8	111	3	US-08-431-184-6	Sequence 6, Appl
55	5	23.8	116	1	US-08-139-862-4	Sequence 4, Appl
56	5	23.8	117	1	US-09-046-479-2	Sequence 2, Appl
57	5	23.8	120	2	US-08-446-345-27	Sequence 27, Appl
58	5	23.8	123	1	US-08-082-623-3	Sequence 3, Appl
59	5	23.8	125	4	US-08-751-359-13	Sequence 13, Appl
60	5	23.8	125	4	US-08-907-146-13	Sequence 13, Appl
61	5	23.8	126	4	US-08-751-359-12	Sequence 12, Appl
62	5	23.8	126	4	US-08-907-146-12	Sequence 12, Appl
63	5	23.8	130	3	US-09-014-574-1	Sequence 1, Appl
64	5	23.8	156	4	US-09-181-183-2	Sequence 2, Appl
65	5	23.8	156	4	US-09-277-700-2	Sequence 2, Appl
66	5	23.8	169	4	US-08-720-625-4	Sequence 4, Appl
67	5	23.8	170	3	US-08-764-563-5	Sequence 5, Appl
68	5	23.8	172	2	US-08-557-309B-39	Sequence 39, Appl
69	5	23.8	172	3	US-08-834-306-39	Sequence 39, Appl
70	5	23.8	172	4	US-08-933-674A-39	Sequence 39, Appl
71	5	23.8	179	3	US-08-764-563-4	Sequence 4, Appl
72	5	23.8	179	4	US-08-862-124-2	Sequence 2, Appl
73	5	23.8	196	2	US-08-685-992-1	Sequence 1, Appl
74	5	23.8	196	2	US-09-144-925-1	Sequence 1, Appl
75	5	23.8	200	6	5189147-8	Sequence 7, Appl
76	5	23.8	212	2	US-08-737-129A-2	Sequence 2, Appl
77	5	23.8	212	5	PCT-US94-14106-59	Sequence 59, Appl
78	5	23.8	219	1	US-08-152-019A-31	Sequence 31, Appl
79	5	23.8	219	1	US-08-152-019A-32	Sequence 32, Appl
80	5	23.8	219	2	US-08-460-309-18	Sequence 18, Appl
81	5	23.8	219	2	US-08-125-077-18	Sequence 18, Appl
82	5	23.8	220	4	US-09-052-089A-3	Sequence 3, Appl
83	5	23.8	222	2	US-08-737-129A-6	Sequence 6, Appl
84	5	23.8	222	2	US-09-181-183-28	Sequence 28, Appl
85	5	23.8	222	4	US-09-277-700-28	Sequence 28, Appl
86	5	23.8	223	5	PCT-US94-14106-51	Sequence 51, Appl
87	5	23.8	223	5	PCT-US94-14106-55	Sequence 55, Appl
88	5	23.8	236	1	US-08-158-682A-4	Sequence 4, Appl
89	5	23.8	236	1	US-08-687-895-4	Sequence 4, Appl
90	5	23.8	236	1	US-08-816-241-4	Sequence 4, Appl
91	5	23.8	236	2	US-09-040-482-4	Sequence 4, Appl
92	5	23.8	236	2	US-09-148-395-4	Sequence 4, Appl
93	5	23.8	236	3	US-08-916-902A-1	Sequence 1, Appl
94	5	23.8	239	2	US-09-213-389-1	Sequence 1, Appl
95	5	23.8	240	2	US-09-194-905-8	Sequence 8, Appl
96	5	23.8	243	3	US-09-286-690-10	Sequence 10, Appl
97	5	23.8	247	3	US-09-100-804-13	Sequence 13, Appl
98	5	23.8	247	5	PCT-US94-07659-2	Sequence 2, Appl
99	5	23.8	254	2	US-08-792-824-4	Sequence 4, Appl
100	5	23.8	254	2	US-08-792-824-4	Sequence 4, Appl

Patent No. 5189147

101	5	23.8	254	2	US-08-792-824-7	Sequence 7, Appl	174	5	23.8	437	2	US-08-989-925-3	Sequence 3, Appl
102	5	23.8	254	2	US-08-792-824-10	Sequence 10, Appl	175	5	23.8	439	4	US-08-716-873-2	Sequence 2, Appl
103	5	23.8	254	2	US-08-792-824-13	Sequence 13, Appl	176	5	23.8	439	4	US-08-716-873-4	Sequence 4, Appl
104	5	23.8	259	1	US-08-203-716-3	Sequence 3, Appl	177	5	23.8	439	4	US-09-368-431-2	Sequence 2, Appl
105	5	23.8	259	1	US-08-440-179-3	Sequence 3, Appl	178	5	23.8	439	4	US-09-368-431-4	Sequence 4, Appl
106	5	23.8	259	2	US-08-432-693-11	Sequence 11, Appl	179	5	23.8	444	4	US-09-252-292C-27	Sequence 27, Appl
107	5	23.8	259	2	US-08-599-895-11	Sequence 11, Appl	180	5	23.8	445	1	US-08-353-400-33	Sequence 33, Appl
108	5	23.8	259	2	US-08-483-806-2	Sequence 2, Appl	181	5	23.8	447	6	5455030-1	Patent No. 5455030
109	5	23.8	259	3	US-09-211-290-11	Sequence 11, Appl	182	5	23.8	450	4	US-09-368-993-41	Sequence 41, Appl
110	5	23.8	259	3	US-09-030-613-11	Sequence 11, Appl	183	5	23.8	464	1	US-08-353-400-36	Sequence 36, Appl
111	5	23.8	259	4	US-09-322-676-11	Sequence 11, Appl	184	5	23.8	464	2	US-08-553-619B-9	Sequence 9, Appl
112	5	23.8	259	4	US-09-039-657-3	Sequence 3, Appl	185	5	23.8	466	4	US-09-268-993-39	Sequence 39, Appl
113	5	23.8	259	4	US-08-748-547-4	Sequence 4, Appl	186	5	23.8	467	4	US-09-086-483A-6	Sequence 6, Appl
114	5	23.8	259	4	US-09-466-036A-11	Sequence 11, Appl	187	5	23.8	467	4	US-08-495-484-12	Sequence 12, Appl
115	5	23.8	259	4	US-09-451-905-11	Sequence 11, Appl	188	5	23.8	468	2	US-08-303-569B-7	Sequence 7, Appl
116	5	23.8	259	5	PC-US91-02339-1	Sequence 1, Appl	189	5	23.8	468	2	US-08-116-247-7	Sequence 7, Appl
117	5	23.8	282	5	PC-US94-09752-3	Sequence 3, Appl	190	5	23.8	468	4	US-09-013-895A-2	Sequence 2, Appl
118	5	23.8	287	1	US-08-457-245-3	Sequence 3, Appl	191	5	23.8	469	2	US-08-968-751-2	Sequence 2, Appl
119	5	23.8	287	4	US-08-862-124-17	Sequence 14, Appl	192	5	23.8	469	4	US-09-052-089A-1	Sequence 1, Appl
120	5	23.8	288	3	US-09-100-804-14	Sequence 14, Appl	193	5	23.8	473	4	US-09-171-945-131	Sequence 131, Appl
121	5	23.8	297	1	US-08-534-910B-7	Sequence 7, Appl	194	5	23.8	477	4	US-09-446-504-19	Sequence 19, Appl
122	5	23.8	297	1	US-08-534-910B-10	Sequence 10, Appl	195	5	23.8	477	4	US-09-712-266-19	Sequence 19, Appl
123	5	23.8	304	4	US-08-862-124-14	Sequence 14, Appl	196	5	23.8	479	1	US-08-484-105-10	Sequence 10, Appl
124	5	23.8	320	2	US-08-579-940-8	Sequence 8, Appl	197	5	23.8	479	1	US-08-484-105-10	Sequence 10, Appl
125	5	23.8	320	2	US-08-579-940-8	Sequence 8, Appl	198	5	23.8	480	2	US-08-828-488-8	Sequence 8, Appl
126	5	23.8	320	3	US-08-564-313-2	Sequence 2, Appl	199	5	23.8	508	2	US-08-724-281-1	Sequence 1, Appl
127	5	23.8	320	4	US-09-069-138-2	Sequence 2, Appl	200	5	23.8	516	4	US-09-201-641-6	Sequence 6, Appl
128	5	23.8	322	4	US-08-964-308-8	Sequence 8, Appl	201	5	23.8	524	4	US-08-624-123-2	Sequence 2, Appl
129	5	23.8	322	4	US-08-964-313-8	Sequence 8, Appl	202	5	23.8	524	1	US-08-624-123-2	Sequence 2, Appl
130	5	23.8	322	4	US-09-069-138-8	Sequence 8, Appl	203	5	23.8	533	2	US-08-770-544-4	Sequence 4, Appl
131	5	23.8	324	2	US-08-579-940-7	Sequence 7, Appl	204	5	23.8	535	3	US-08-813-574-2	Sequence 2, Appl
132	5	23.8	328	2	US-08-977-767-1	Sequence 1, Appl	205	5	23.8	538	2	US-08-541-759B-2	Sequence 2, Appl
133	5	23.8	345	1	US-08-446-345-40	Sequence 40, Appl	206	5	23.8	580	1	US-08-309-513-6	Sequence 6, Appl
134	5	23.8	361	1	US-08-537-434-1	Sequence 1, Appl	207	5	23.8	580	1	PCT-US92-08756A-6	Sequence 6, Appl
135	5	23.8	363	1	US-08-484-105-20	Sequence 20, Appl	208	5	23.8	585	1	US-08-477-674-10	Sequence 10, Appl
136	5	23.8	363	1	US-08-484-105-20	Sequence 20, Appl	209	5	23.8	585	1	US-08-473-791-10	Sequence 10, Appl
137	5	23.8	363	2	US-08-646-590B-36	Sequence 36, Appl	210	5	23.8	585	2	US-08-316-714-10	Sequence 10, Appl
138	5	23.8	363	4	US-09-412-184-36	Sequence 36, Appl	211	5	23.8	585	3	US-08-473-673-10	Sequence 10, Appl
139	5	23.8	369	1	US-08-700-359-21	Sequence 21, Appl	212	5	23.8	585	3	US-08-442-542-18	Sequence 18, Appl
140	5	23.8	370	3	US-08-911-321-10	Sequence 10, Appl	213	5	23.8	599	3	US-08-765-469-18	Sequence 18, Appl
141	5	23.8	374	3	US-08-609-236-6	Sequence 6, Appl	214	5	23.8	626	2	US-08-596-300A-7	Sequence 7, Appl
142	5	23.8	375	2	US-08-494-151-14	Sequence 14, Appl	215	5	23.8	626	2	US-08-596-300A-14	Sequence 14, Appl
143	5	23.8	375	3	US-09-106-217-16	Sequence 16, Appl	216	5	23.8	630	4	US-09-704-449-2	Sequence 2, Appl
144	5	23.8	375	4	US-09-171-337A-7	Sequence 7, Appl	217	5	23.8	638	4	US-09-228-986-74	Sequence 74, Appl
145	5	23.8	375	4	US-08-588-113-2	Sequence 8, Appl	218	5	23.8	639	1	US-08-466-390-2	Sequence 2, Appl
146	5	23.8	376	1	US-08-588-113-2	Sequence 2, Appl	219	5	23.8	639	1	US-08-470-950-2	Sequence 2, Appl
147	5	23.8	377	3	US-09-106-217-2	Sequence 2, Appl	220	5	23.8	639	1	US-08-467-781-2	Sequence 2, Appl
148	5	23.8	379	3	US-08-840-204-3	Sequence 3, Appl	221	5	23.8	639	1	US-08-195-487-2	Sequence 2, Appl
149	5	23.8	379	4	US-09-193-266-1	Sequence 1, Appl	222	5	23.8	639	5	PCT-US93-06160-2	Sequence 2, Appl
150	5	23.8	385	1	US-08-036-210-5	Sequence 5, Appl	223	5	23.8	639	5	US-09-136-282-2	Sequence 2, Appl
151	5	23.8	385	2	US-08-449-609-5	Sequence 5, Appl	224	5	23.8	643	4	US-09-178-252-25	Sequence 25, Appl
152	5	23.8	390	1	US-08-121-714-6	Sequence 6, Appl	225	5	23.8	647	4	US-09-423-439-60	Sequence 60, Appl
153	5	23.8	390	1	US-08-477-108A-6	Sequence 6, Appl	226	5	23.8	670	4	US-08-980-080-2	Sequence 2, Appl
154	5	23.8	390	2	US-08-477-113-6	Sequence 6, Appl	227	5	23.8	685	2	US-08-878-989-1	Sequence 1, Appl
155	5	23.8	390	2	PC-US93-08322-6	Sequence 6, Appl	228	5	23.8	685	3	US-09-136-282-2	Sequence 2, Appl
156	5	23.8	392	2	US-08-886-152-3	Sequence 1, Appl	229	5	23.8	685	4	US-09-272-796-1	Sequence 1, Appl
157	5	23.8	392	2	US-08-886-152-3	Sequence 1, Appl	230	5	23.8	685	4	US-09-505-744-2	Sequence 2, Appl
158	5	23.8	394	4	US-09-144-914-4	Sequence 4, Appl	231	5	23.8	713	4	US-09-346-237-9	Sequence 9, Appl
159	5	23.8	397	3	US-09-079-415-2	Sequence 2, Appl	232	5	23.8	726	3	US-09-129-075-4	Sequence 4, Appl
160	5	23.8	398	1	US-08-328-314-2	Sequence 2, Appl	233	5	23.8	726	4	US-09-346-237-3	Sequence 3, Appl
161	5	23.8	398	1	US-08-731-045-2	Sequence 2, Appl	234	5	23.8	726	4	US-09-346-237-13	Sequence 13, Appl
162	5	23.8	402	1	US-08-315-461-7	Sequence 7, Appl	235	5	23.8	756	1	US-08-434-730-16	Sequence 16, Appl
163	5	23.8	402	3	US-08-948-997-4	Sequence 4, Appl	236	5	23.8	767	4	US-08-836-567-8	Sequence 8, Appl
164	5	23.8	402	3	US-08-840-204-2	Sequence 2, Appl	237	5	23.8	774	1	US-07-731-157A-7	Sequence 7, Appl
165	5	23.8	402	4	US-09-348-817A-4	Sequence 4, Appl	238	5	23.8	774	2	US-08-541-780-7	Sequence 7, Appl
166	5	23.8	402	4	US-09-026-408-3	Sequence 3, Appl	239	5	23.8	799	3	US-08-909-954-4	Sequence 4, Appl
167	5	23.8	412	1	US-08-208-007A-12	Sequence 12, Appl	240	5	23.8	820	2	US-08-380-182-23	Sequence 23, Appl
168	5	23.8	412	4	US-08-974-691-4	Sequence 4, Appl	241	5	23.8	835	4	US-09-291-839-2	Sequence 2, Appl
169	5	23.8	415	4	US-09-100-193-2	Sequence 2, Appl	242	5	23.8	859	1	US-08-395-580-2	Sequence 2, Appl
170	5	23.8	417	3	US-08-705-771-18	Sequence 18, Appl	243	5	23.8	859	1	PCT-US95-02752-2	Sequence 2, Appl
171	5	23.8	436	4	US-08-716-873-5	Sequence 5, Appl	244	5	23.8	913	3	US-08-911-853-17	Sequence 17, Appl
172	5	23.8	436	4	US-09-368-431-5	Sequence 5, Appl	245	5	23.8	913	3	US-09-479-409-17	Sequence 17, Appl
173	5	23.8	437	1	US-08-764-343-1	Sequence 1, Appl	246	5	23.8	913	4	US-09-479-453-17	Sequence 17, Appl

247	5	23.8	948	1	US-08-698-551-14	Sequence 14, Appl	320	4	19.0	7	4	US-08-868-452-8	Sequence 8, Appl
248	5	23.8	948	2	US-08-602-228-14	Sequence 14, Appl	321	4	19.0	7	6	5204326-110	Patent No. 5204326
249	5	23.8	948	3	US-08-533-901B-14	Sequence 14, Appl	322	4	19.0	8	2	US-08-846-360-141	Sequence 141, Appl
250	5	23.8	948	4	US-08-839-032A-14	Sequence 14, Appl	323	4	19.0	8	2	US-08-792-553-13	Sequence 13, Appl
251	5	23.8	948	5	US-08-839-032A-14	Sequence 14, Appl	324	4	19.0	8	4	US-09-136-389-141	Sequence 141, Appl
252	5	23.8	948	6	PCT-US95-1272A-14	Sequence 14, Appl	325	4	19.0	8	6	5449756-6	Patent No. 5449756
253	5	23.8	956	1	US-08-897-443-3	Sequence 3, Appl	326	4	19.0	9	1	US-08-619-645-1	Sequence 1, Appl
254	5	23.8	966	1	US-08-571-758-2	Sequence 2, Appl	327	4	19.0	9	2	US-08-634-493-1	Sequence 1, Appl
255	5	23.8	966	2	US-08-909-984A-2	Sequence 2, Appl	328	4	19.0	9	3	US-09-079-447-1	Sequence 1, Appl
256	5	23.8	1050	1	US-08-909-983-2	Sequence 2, Appl	329	4	19.0	9	4	US-08-660-092-72	Sequence 72, Appl
257	5	23.8	1050	2	US-09-045-632-49	Sequence 49, Appl	330	4	19.0	10	1	US-07-864-475A-11	Sequence 11, Appl
258	5	23.8	1050	3	US-09-045-632-50	Sequence 50, Appl	331	4	19.0	10	1	US-08-218-026-38	Sequence 38, Appl
259	5	23.8	1066	4	US-09-004-838-24	Sequence 24, Appl	332	4	19.0	10	2	US-08-724-548-9	Sequence 9, Appl
260	5	23.8	1141	1	US-08-131-365B-54	Sequence 54, Appl	333	4	19.0	10	2	US-08-724-548-10	Sequence 10, Appl
261	5	23.8	1141	2	US-08-668-123-54	Sequence 54, Appl	334	4	19.0	10	2	US-08-724-548-11	Sequence 11, Appl
262	5	23.8	1186	4	US-09-178-252-23	Sequence 23, Appl	335	4	19.0	10	2	US-08-724-548-12	Sequence 12, Appl
263	5	23.8	1203	4	US-09-075-272-4	Sequence 4, Appl	336	4	19.0	10	2	US-08-724-548-13	Sequence 13, Appl
264	5	23.8	1222	4	US-09-004-838-119	Sequence 119, Appl	337	4	19.0	10	2	US-08-724-548-14	Sequence 14, Appl
265	5	23.8	1227	1	US-08-448-170-8	Sequence 8, Appl	338	4	19.0	10	2	US-08-724-548-15	Sequence 15, Appl
266	5	23.8	1227	2	US-08-961-803-9	Sequence 9, Appl	339	4	19.0	10	2	US-08-653-632-38	Sequence 38, Appl
267	5	23.8	1275	3	US-09-120-513-2	Sequence 2, Appl	340	4	19.0	10	2	US-08-637-759B-192	Sequence 192, Appl
268	5	23.8	1275	4	US-09-450-105-2	Sequence 2, Appl	341	4	19.0	10	2	US-08-468-249A-11	Sequence 11, Appl
269	5	23.8	1279	2	US-08-784-649A-2	Sequence 2, Appl	342	4	19.0	10	3	US-08-596-257A-8	Sequence 8, Appl
270	5	23.8	1280	2	US-08-583-276-19	Sequence 19, Appl	343	4	19.0	10	3	US-08-871-355A-192	Sequence 192, Appl
271	5	23.8	1280	3	US-08-752-447-2	Sequence 2, Appl	344	4	19.0	10	3	US-07-978-674B-9	Sequence 9, Appl
272	5	23.8	1283	6	5206352-4	Patent No. 5206352	345	4	19.0	10	6	5204326-111	Patent No. 5204326
273	5	23.8	1323	4	US-09-004-838-90	Sequence 90, Appl	346	4	19.0	10	3	US-07-978-674B-10	Sequence 10, Appl
274	5	23.8	1434	2	US-08-540-406-10	Sequence 10, Appl	347	4	19.0	10	3	US-07-978-674B-11	Sequence 11, Appl
275	5	23.8	1434	3	US-08-656-055-10	Sequence 10, Appl	348	4	19.0	10	3	US-07-978-674B-12	Sequence 12, Appl
276	5	23.8	1434	4	US-08-954-668-10	Sequence 10, Appl	349	4	19.0	10	3	US-07-978-674B-13	Sequence 13, Appl
277	5	23.8	1434	5	PCT-US95-13233-10	Sequence 10, Appl	350	4	19.0	10	3	US-07-978-674B-14	Sequence 14, Appl
278	5	23.8	1581	4	US-09-110-517-2	Sequence 2, Appl	351	4	19.0	10	3	US-08-860-339-8	Sequence 8, Appl
279	5	23.8	1786	2	US-08-477-451-16	Sequence 16, Appl	352	4	19.0	10	3	US-08-939-853A-16	Sequence 16, Appl
280	5	23.8	1786	4	US-08-973-462-8	Sequence 8, Appl	353	4	19.0	10	4	US-09-201-945-192	Sequence 192, Appl
281	5	23.8	1817	4	US-09-004-838-125	Sequence 125, Appl	354	4	19.0	10	4	5204326-111	Patent No. 5204326
282	5	23.8	1872	6	5386025-6	Patent No. 5386025	355	4	19.0	12	1	US-08-309-512-56	Sequence 56, Appl
283	5	23.8	1873	1	US-08-435-675B-4	Sequence 4, Appl	356	4	19.0	12	2	US-08-479-614-10	Sequence 10, Appl
284	5	23.8	1873	2	US-08-336-257A-7	Sequence 7, Appl	357	4	19.0	12	2	US-08-209-525-38	Sequence 38, Appl
285	5	23.8	1940	2	US-08-644-271-30	Sequence 30, Appl	358	4	19.0	13	1	US-08-619-645-5	Sequence 5, Appl
286	5	23.8	2441	3	US-08-194-468-2	Sequence 2, Appl	359	4	19.0	13	1	US-08-619-645-6	Sequence 6, Appl
287	5	23.8	2441	4	US-08-961-739-2	Sequence 2, Appl	360	4	19.0	13	2	US-08-634-493-5	Sequence 5, Appl
288	5	23.8	2516	3	US-08-374-077C-2	Sequence 2, Appl	361	4	19.0	13	2	US-08-634-493-6	Sequence 6, Appl
289	5	23.8	2516	4	US-08-895-590-2	Sequence 2, Appl	362	4	19.0	13	4	US-08-423-646A-54	Sequence 54, Appl
290	5	23.8	2802	4	US-09-542-331-1	Sequence 1, Appl	363	4	19.0	13	5	PCT-US93-05647-11	Sequence 11, Appl
291	5	23.8	3072	4	US-09-413-814-93	Sequence 93, Appl	364	4	19.0	15	1	US-08-268-251-27	Sequence 27, Appl
292	5	23.8	3079	4	US-09-413-814-80	Sequence 80, Appl	365	4	19.0	15	5	PCT-US93-01112-27	Sequence 27, Appl
293	5	23.8	3118	2	US-08-457-273B-8	Sequence 8, Appl	366	4	19.0	16	1	US-07-940-861-7	Sequence 7, Appl
294	5	23.8	3119	1	US-08-246-982A-16	Sequence 16, Appl	367	4	19.0	16	1	US-08-459-512-7	Sequence 7, Appl
295	5	23.8	3119	2	US-08-453-265-16	Sequence 16, Appl	368	4	19.0	16	2	US-08-459-512-7	Sequence 7, Appl
296	5	23.8	3144	1	US-08-246-982A-6	Sequence 6, Appl	369	4	19.0	16	2	US-08-460-132-7	Sequence 7, Appl
297	5	23.8	3144	2	US-08-453-265-6	Sequence 6, Appl	370	4	19.0	16	2	US-09-133-774-5	Sequence 5, Appl
298	5	23.8	3144	3	US-08-457-273B-42	Sequence 42, Appl	371	4	19.0	16	2	US-08-312-202B-4	Sequence 4, Appl
299	5	23.8	3144	4	US-08-556-419-21	Sequence 21, Appl	372	4	19.0	16	3	US-09-303-862-5	Sequence 5, Appl
300	5	23.8	3144	5	US-09-041-886-15	Sequence 15, Appl	373	4	19.0	16	3	US-09-079-347-4	Sequence 4, Appl
301	5	23.8	3257	3	US-09-335-409-5	Sequence 5, Appl	374	4	19.0	16	3	US-09-075-725-4	Sequence 4, Appl
302	5	23.8	3257	4	US-09-568-102-5	Sequence 5, Appl	375	4	19.0	16	4	US-09-171-705-40	Sequence 40, Appl
303	5	23.8	3257	5	US-09-567-969-5	Sequence 5, Appl	376	4	19.0	16	4	US-09-171-705-41	Sequence 41, Appl
304	5	23.8	3257	6	US-09-568-480-5	Sequence 5, Appl	377	4	19.0	16	4	US-08-809-646-4	Sequence 4, Appl
305	5	23.8	3257	7	US-09-568-486-5	Sequence 5, Appl	378	4	19.0	16	5	PCT-US92-02050-7	Sequence 7, Appl
306	5	23.8	3257	8	US-09-568-472-5	Sequence 5, Appl	379	4	19.0	17	2	PCT-US95-12433-4	Sequence 4, Appl
307	5	19.0	5	1	US-08-477-727A-54	Sequence 54, Appl	380	4	19.0	17	6	US-08-850-910A-15	Sequence 15, Appl
308	4	19.0	6	2	US-08-724-548-50	Sequence 50, Appl	381	4	19.0	17	6	5185441-3	Patent No. 5185441
309	4	19.0	6	3	US-08-724-548-51	Sequence 51, Appl	382	4	19.0	18	1	US-08-476-405A-1	Sequence 1, Appl
310	4	19.0	6	4	US-07-978-674B-50	Sequence 50, Appl	383	4	19.0	18	1	US-08-499-523-15	Sequence 15, Appl
311	4	19.0	6	5	US-07-978-674B-51	Sequence 51, Appl	384	4	19.0	18	4	US-09-128-345-15	Sequence 15, Appl
312	4	19.0	6	6	US-09-156-580-6	Sequence 6, Appl	385	4	19.0	19	1	US-08-131-365B-44	Sequence 44, Appl
313	4	19.0	6	7	5204326-147	Patent No. 5204326	386	4	19.0	19	2	US-08-668-123-44	Sequence 44, Appl
314	4	19.0	6	8	US-08-470-837-8	Sequence 8, Appl	387	4	19.0	20	1	US-07-678-974D-19	Sequence 19, Appl
315	4	19.0	7	1	US-08-632-598-25	Sequence 25, Appl	388	4	19.0	20	1	US-08-476-405A-2	Sequence 2, Appl
316	4	19.0	7	2	US-09-173-941-16	Sequence 16, Appl	389	4	19.0	20	2	US-08-850-910A-34	Sequence 34, Appl
317	4	19.0	7	3	US-09-173-941-16	Sequence 16, Appl	390	4	19.0	20	2	US-08-945-168-24	Sequence 24, Appl
318	4	19.0	7	4	US-09-173-941-100	Sequence 100, Appl	391	4	19.0	20	4	US-08-737-226-11	Sequence 11, Appl
319	4	19.0	7	5	US-09-231-240-25	Sequence 25, Appl	392	4	19.0	20	4	US-09-324-910-2	Sequence 2, Appl

393	4	19.0	21	1	US-07-851-976B-1	Sequence 1, Appli	466	4	19.0	31	6	5204326-142	Patent No. 5204326
394	4	19.0	21	1	US-07-508-455A-7	Sequence 7, Appli	467	4	19.0	31	6	5204326-144	Patent No. 5204326
395	4	19.0	21	1	US-08-434-120-33	Sequence 33, Appli	468	4	19.0	32	1	US-07-754-958-5	Sequence 5, Appli
396	4	19.0	21	1	US-08-291-609-1	Sequence 1, Appli	469	4	19.0	32	1	US-07-754-947-5	Sequence 5, Appli
397	4	19.0	21	1	US-08-465-325-33	Sequence 33, Appli	470	4	19.0	32	1	US-07-952-735A-3	Sequence 3, Appli
398	4	19.0	21	1	US-08-401-136-1	Sequence 1, Appli	471	4	19.0	32	1	US-07-952-735A-4	Sequence 4, Appli
399	4	19.0	21	1	US-08-850-554-1	Sequence 1, Appli	472	4	19.0	32	1	US-08-305-799A-1	Sequence 1, Appli
400	4	19.0	21	3	US-09-115-737-33	Sequence 33, Appli	473	4	19.0	32	2	US-08-526-583-5	Sequence 5, Appli
401	4	19.0	21	6	5182195-5	Patent No. 5182195	474	4	19.0	32	2	US-08-526-583-6	Sequence 6, Appli
402	4	19.0	22	3	US-08-833-553-6	Sequence 6, Appli	475	4	19.0	32	2	US-08-526-583-7	Sequence 7, Appli
403	4	19.0	22	3	US-09-418-222-6	Sequence 6, Appli	476	4	19.0	32	2	US-08-526-583-8	Sequence 8, Appli
404	4	19.0	22	4	US-08-706-344C-8	Sequence 8, Appli	477	4	19.0	32	2	US-08-526-583C-41	Sequence 41, Appli
405	4	19.0	22	4	US-09-439-313-331	Sequence 331, App	478	4	19.0	32	3	US-08-954-915A-44	Sequence 44, Appli
406	4	19.0	23	1	US-07-656-566-1	Sequence 1, Appli	479	4	19.0	32	3	US-08-954-915A-45	Sequence 45, Appli
407	4	19.0	23	1	US-08-168-251-57	Sequence 57, Appli	480	4	19.0	32	3	US-08-484-322-67	Sequence 67, Appli
408	4	19.0	23	1	US-08-146-028-28	Sequence 28, Appli	481	4	19.0	32	3	US-08-484-322-67	Sequence 10, Appli
409	4	19.0	23	2	US-08-146-028-167	Sequence 167, App	482	4	19.0	32	3	US-07-401-432-10	Sequence 50, Appli
410	4	19.0	23	4	US-08-723-425A-28	Sequence 28, Appli	483	4	19.0	32	3	US-07-401-432-50	Sequence 51, Appli
411	4	19.0	23	4	US-08-723-425A-167	Sequence 167, App	484	4	19.0	32	4	US-07-401-432-51	Sequence 415, App
412	4	19.0	23	4	US-09-112-206-28	Sequence 28, Appli	485	4	19.0	32	4	US-08-905-223-415	Sequence 415, App
413	4	19.0	23	4	US-09-112-206-167	Sequence 167, App	486	4	19.0	32	6	5183802-5	Patent No. 5183802
414	4	19.0	23	5	PCT-US93-01112-57	Sequence 57, Appli	487	4	19.0	32	6	5204326-143	Patent No. 5204326
415	4	19.0	24	3	US-08-701-382-8	Sequence 8, Appli	488	4	19.0	32	6	5208144-17	Patent No. 5208144
416	4	19.0	24	3	US-08-894-699-35	Sequence 35, Appli	489	4	19.0	33	1	US-07-776-272-9	Sequence 9, Appli
417	4	19.0	24	4	US-09-444-410-35	Sequence 35, Appli	490	4	19.0	33	2	US-08-078-311-26	Sequence 26, Appli
418	4	19.0	24	4	US-08-788-820-8	Sequence 8, Appli	491	4	19.0	33	2	US-08-460-402-26	Sequence 26, Appli
419	4	19.0	25	2	US-08-632-598-43	Sequence 43, Appli	492	4	19.0	33	4	US-09-066-046-9	Sequence 9, Appli
420	4	19.0	25	2	US-08-632-598-44	Sequence 44, Appli	493	4	19.0	34	1	US-07-952-735A-1	Sequence 1, Appli
421	4	19.0	25	2	US-08-632-598-45	Sequence 45, Appli	494	4	19.0	34	1	US-07-952-735A-2	Sequence 2, Appli
422	4	19.0	25	2	US-08-632-598-46	Sequence 46, Appli	495	4	19.0	34	4	US-08-974-549A-38	Sequence 38, Appli
423	4	19.0	25	2	US-08-765-452-15	Sequence 15, Appli	496	4	19.0	34	4	US-08-482-918-79	Sequence 79, Appli
424	4	19.0	25	4	US-09-231-240-43	Sequence 43, Appli	497	4	19.0	34	4	US-09-224-681-79	Sequence 79, Appli
425	4	19.0	25	4	US-09-231-240-44	Sequence 44, Appli	498	4	19.0	34	4	US-08-336-728A-79	Sequence 79, Appli
426	4	19.0	25	4	US-09-231-240-45	Sequence 45, Appli	499	4	19.0	34	6	5204326-2	Patent No. 5204326
427	4	19.0	25	4	US-09-231-240-46	Sequence 46, Appli	500	4	19.0	35	4	US-09-227-357-183	Sequence 183, App
428	4	19.0	26	1	US-07-778-847-2	Sequence 2, Appli	501	4	19.0	36	1	US-08-487-890A-32	Sequence 32, Appli
429	4	19.0	26	2	US-08-856-006-2	Sequence 2, Appli	502	4	19.0	36	2	US-08-478-435-32	Sequence 32, Appli
430	4	19.0	26	2	US-08-620-151-32	Sequence 32, Appli	503	4	19.0	36	2	US-08-337-483-32	Sequence 32, Appli
431	4	19.0	26	2	US-08-850-910A-3	Sequence 3, Appli	504	4	19.0	36	2	US-08-478-373-32	Sequence 32, Appli
432	4	19.0	26	3	US-09-122-642-2	Sequence 2, Appli	505	4	19.0	36	3	US-08-474-671-32	Sequence 32, Appli
433	4	19.0	26	3	US-08-851-843A-145	Sequence 145, App	506	4	19.0	36	3	US-08-483-577A-32	Sequence 32, Appli
434	4	19.0	26	4	US-08-974-549A-263	Sequence 263, App	507	4	19.0	36	4	US-08-897-438-32	Sequence 32, Appli
435	4	19.0	26	4	US-08-854-050-145	Sequence 145, App	508	4	19.0	36	4	US-08-637-654-32	Sequence 32, Appli
436	4	19.0	26	4	US-09-424-014-2	Sequence 2, Appli	509	4	19.0	37	1	US-08-446-692-15	Sequence 15, Appli
437	4	19.0	26	4	US-09-430-323-145	Sequence 145, App	510	4	19.0	37	2	US-08-488-351A-15	Sequence 15, Appli
438	4	19.0	27	1	US-08-446-692-7	Sequence 7, Appli	511	4	19.0	37	4	US-08-482-918-80	Sequence 80, Appli
439	4	19.0	27	2	US-08-488-351A-7	Sequence 6, Appli	512	4	19.0	37	4	US-09-224-681-80	Sequence 80, Appli
440	4	19.0	27	3	US-08-749-816-6	Sequence 6, Appli	513	4	19.0	37	4	US-08-336-728A-80	Sequence 80, Appli
441	4	19.0	27	3	US-08-318-794-29	Sequence 29, Appli	514	4	19.0	38	1	US-08-176-500-46	Sequence 46, Appli
442	4	19.0	27	3	US-09-100-409A-43	Sequence 43, Appli	515	4	19.0	38	1	US-08-471-052A-46	Sequence 46, Appli
443	4	19.0	27	3	US-08-851-843A-150	Sequence 150, App	516	4	19.0	38	2	US-08-189-331-46	Sequence 46, Appli
444	4	19.0	27	4	US-08-974-549A-270	Sequence 270, App	517	4	19.0	38	2	US-08-471-939-46	Sequence 46, Appli
445	4	19.0	27	4	US-08-854-050-150	Sequence 150, App	518	4	19.0	38	2	US-08-471-800-46	Sequence 46, Appli
446	4	19.0	27	4	US-09-144-914-10	Sequence 10, Appli	519	4	19.0	38	2	US-08-471-068-46	Sequence 46, Appli
447	4	19.0	27	4	US-09-430-323-150	Sequence 150, App	520	4	19.0	39	1	US-08-324-301-6	Sequence 6, Appli
448	4	19.0	27	4	US-08-470-106-29	Sequence 29, Appli	521	4	19.0	39	3	US-08-258-287B-46	Sequence 46, Appli
449	4	19.0	27	5	PCT-US95-13841-10	Sequence 10, Appli	522	4	19.0	39	3	US-08-368-704C-45	Sequence 45, Appli
450	4	19.0	29	1	US-07-754-958-6	Sequence 6, Appli	523	4	19.0	40	1	US-08-188-228-12	Sequence 12, Appli
451	4	19.0	29	1	US-07-754-947-6	Sequence 6, Appli	524	4	19.0	40	1	US-08-188-228-18	Sequence 18, Appli
452	4	19.0	29	4	US-09-227-357-402	Sequence 402, App	525	4	19.0	40	1	US-08-188-228-24	Sequence 24, Appli
453	4	19.0	30	1	US-08-305-799A-10	Sequence 10, Appli	526	4	19.0	40	1	US-08-332-643-12	Sequence 12, Appli
454	4	19.0	30	2	US-08-753-829A-13	Sequence 13, Appli	527	4	19.0	40	1	US-08-332-643-24	Sequence 24, Appli
455	4	19.0	30	4	US-08-945-168-107	Sequence 107, App	528	4	19.0	40	1	US-08-332-638-12	Sequence 12, Appli
456	4	19.0	30	4	US-09-230-405-11	Sequence 11, Appli	529	4	19.0	40	1	US-08-332-638-18	Sequence 18, Appli
457	4	19.0	31	1	US-08-190-802A-236	Sequence 236, App	530	4	19.0	40	1	US-08-332-638-24	Sequence 24, Appli
458	4	19.0	31	1	US-08-732-751-3	Sequence 3, Appli	531	4	19.0	40	1	US-08-031-538-52	Sequence 52, Appli
459	4	19.0	31	1	US-08-732-751-4	Sequence 4, Appli	532	4	19.0	40	2	US-08-968-685A-12	Sequence 12, Appli
460	4	19.0	31	3	US-07-401-432-49	Sequence 49, Appli	533	4	19.0	40	4	US-08-988-856B-19	Sequence 19, Appli
461	4	19.0	31	4	US-08-477-346-236	Sequence 236, App	534	4	19.0	40	4	US-08-630-915A-83	Sequence 83, Appli
462	4	19.0	31	4	US-08-473-089-236	Sequence 236, App	535	4	19.0	40	4	US-08-984-277-8	Sequence 8, Appli
463	4	19.0	31	6	5204326-138	Patent No. 5204326	536	4	19.0	41	3	US-08-984-277-8	Sequence 11, Appli
464	4	19.0	31	6	5204326-139	Patent No. 5204326	537	4	19.0	41	4	US-09-156-580-11	Sequence 11, Appli
465	4	19.0	31	6	5204326-141	Patent No. 5204326	538	4	19.0	41	4	US-08-988-856B-18	Sequence 18, Appli

539	4	19.0	4	41	4	US-09-156-579C-13	Sequence 13, Appl	612	4	19.0	64	2	US-08-765-179B-23	Sequence 23, Appl
540	4	19.0	4	42	2	US-08-476-866-23	Sequence 23, Appl	613	4	19.0	64	3	US-08-480-640A-194	Sequence 194, App
541	4	19.0	4	43	4	US-09-156-580-10	Sequence 10, Appl	614	4	19.0	64	4	US-08-686-968C-194	Sequence 194, App
542	4	19.0	4	44	4	US-09-156-579C-12	Sequence 12, Appl	615	4	19.0	64	4	US-08-488-237A-194	Sequence 194, App
543	4	19.0	4	45	1	US-08-325-547-6	Sequence 6, Appl	616	4	19.0	64	4	US-08-375-992A-194	Sequence 194, App
544	4	19.0	4	46	4	US-08-968-685A-1	Sequence 1, Appl	617	4	19.0	65	4	US-09-197-636-6	Sequence 6, Appl
545	4	19.0	4	47	4	US-09-156-580-9	Sequence 9, Appl	618	4	19.0	65	4	US-09-461-697-447	Sequence 447, App
546	4	19.0	4	48	4	US-09-156-580-12	Sequence 12, Appl	619	4	19.0	65	4	US-09-235-451-9	Sequence 9, Appl
547	4	19.0	4	49	4	US-09-156-580-13	Sequence 13, Appl	620	4	19.0	66	1	US-07-626-618A-16	Sequence 16, Appl
548	4	19.0	4	50	4	US-09-156-579C-11	Sequence 11, Appl	621	4	19.0	66	1	US-07-928-611-16	Sequence 16, Appl
549	4	19.0	4	51	4	US-09-156-579C-14	Sequence 14, Appl	622	4	19.0	66	1	US-08-333-977-16	Sequence 16, Appl
550	4	19.0	4	52	4	US-09-156-579C-15	Sequence 15, Appl	623	4	19.0	66	2	US-08-487-811A-16	Sequence 16, Appl
551	4	19.0	4	53	4	PCT-US92-00282-21	Sequence 21, Appl	624	4	19.0	66	4	US-09-060-694-16	Sequence 16, Appl
552	4	19.0	4	54	5	US-07-689-693B-20	Sequence 20, Appl	625	4	19.0	66	5	PCT-US93-07370-16	Sequence 16, Appl
553	4	19.0	4	55	2	US-08-935-450-11	Sequence 11, Appl	626	4	19.0	67	4	US-08-905-223-379	Sequence 379, App
554	4	19.0	4	56	4	US-09-257-218-6	Sequence 6, Appl	627	4	19.0	67	4	US-08-905-223-435	Sequence 435, App
555	4	19.0	4	57	4	US-09-311-760-6	Sequence 6, Appl	628	4	19.0	67	4	US-09-461-697-445	Sequence 445, App
556	4	19.0	4	58	1	US-08-222-616-10	Sequence 10, Appl	629	4	19.0	67	4	US-09-243-675-4	Sequence 4, Appl
557	4	19.0	4	59	4	US-08-939-853A-15	Sequence 15, Appl	630	4	19.0	68	3	US-08-833-553-9	Sequence 9, Appl
558	4	19.0	4	60	4	US-08-963-851-12	Sequence 12, Appl	631	4	19.0	68	4	US-09-418-222-9	Sequence 9, Appl
559	4	19.0	4	61	4	PCT-US95-04228-10	Sequence 10, Appl	632	4	19.0	68	4	US-08-963-851-15	Sequence 15, Appl
560	4	19.0	4	62	5	US-09-222-373-5	Sequence 5, Appl	633	4	19.0	69	1	US-08-056-200-115	Sequence 115, App
561	4	19.0	4	63	3	US-09-222-373-7	Sequence 7, Appl	634	4	19.0	69	2	US-08-800-644-115	Sequence 115, App
562	4	19.0	4	64	4	US-08-996-679-59	Sequence 59, Appl	635	4	19.0	69	3	US-08-851-843A-13	Sequence 13, Appl
563	4	19.0	4	65	4	US-08-945-997C-16	Sequence 16, Appl	636	4	19.0	69	4	US-08-854-050-13	Sequence 13, Appl
564	4	19.0	4	66	4	US-08-945-997C-18	Sequence 18, Appl	637	4	19.0	70	1	US-09-430-323-13	Sequence 13, Appl
565	4	19.0	4	67	4	US-08-945-997C-19	Sequence 19, Appl	638	4	19.0	70	1	US-08-691-641-15	Sequence 15, Appl
566	4	19.0	4	68	4	US-09-001-511-5	Sequence 5, Appl	639	4	19.0	70	2	US-08-724-548-54	Sequence 54, Appl
567	4	19.0	4	69	4	US-09-001-511-7	Sequence 7, Appl	640	4	19.0	70	3	US-07-978-674B-54	Sequence 54, Appl
568	4	19.0	4	70	4	US-09-113-977C-69	Sequence 69, Appl	641	4	19.0	73	1	US-07-689-693B-3	Sequence 3, Appl
569	4	19.0	4	71	4	US-09-510-616-5	Sequence 5, Appl	642	4	19.0	73	1	US-08-624-123-12	Sequence 12, Appl
570	4	19.0	4	72	4	US-09-510-616-7	Sequence 7, Appl	643	4	19.0	73	2	US-08-518-967-1	Sequence 1, Appl
571	4	19.0	4	73	4	US-09-248-074-78	Sequence 78, Appl	644	4	19.0	73	5	PCT-US96-05262-13	Sequence 13, Appl
572	4	19.0	4	74	4	US-09-187-859-62	Sequence 62, Appl	645	4	19.0	76	5	PCT-US95-13658-6	Sequence 6, Appl
573	4	19.0	4	75	3	US-09-222-373-8	Sequence 8, Appl	646	4	19.0	77	1	US-08-014-153D-40	Sequence 40, Appl
574	4	19.0	4	76	4	US-09-001-511-8	Sequence 8, Appl	647	4	19.0	77	2	US-08-570-227A-8	Sequence 8, Appl
575	4	19.0	4	77	4	US-09-510-616-8	Sequence 8, Appl	648	4	19.0	77	3	US-09-077-991-10	Sequence 10, Appl
576	4	19.0	4	78	3	US-09-222-373-46	Sequence 46, Appl	649	4	19.0	82	1	US-08-476-405A-25	Sequence 25, Appl
577	4	19.0	4	79	4	US-09-174-465D-10	Sequence 10, Appl	650	4	19.0	82	2	US-09-382-155-7	Sequence 7, Appl
578	4	19.0	4	80	4	US-09-001-511-46	Sequence 46, Appl	651	4	19.0	82	4	US-08-482-918-59	Sequence 59, Appl
579	4	19.0	4	81	4	US-09-510-616-46	Sequence 46, Appl	652	4	19.0	82	4	US-09-224-681-59	Sequence 59, Appl
580	4	19.0	4	82	3	US-08-858-207A-289	Sequence 289, App	653	4	19.0	82	4	US-09-074-044A-7	Sequence 7, Appl
581	4	19.0	4	83	5	US-08-630-916A-79	Sequence 79, Appl	654	4	19.0	82	4	US-08-336-728A-59	Sequence 59, Appl
582	4	19.0	4	84	5	US-08-905-223-326	Sequence 326, App	655	4	19.0	85	3	US-08-236-886-7	Sequence 7, Appl
583	4	19.0	4	85	2	US-08-706-741B-52	Sequence 52, Appl	656	4	19.0	85	4	US-09-203-623-4	Sequence 4, Appl
584	4	19.0	4	86	1	US-08-924-695A-52	Sequence 52, Appl	657	4	19.0	86	2	US-08-053-451B-126	Sequence 126, App
585	4	19.0	4	87	2	US-08-485-455D-73	Sequence 73, Appl	658	4	19.0	86	4	US-09-382-155-8	Sequence 8, Appl
586	4	19.0	4	88	2	US-08-592-406-22	Sequence 22, Appl	659	4	19.0	86	4	US-09-074-044A-8	Sequence 8, Appl
587	4	19.0	4	89	2	US-08-482-130C-73	Sequence 73, Appl	660	4	19.0	87	2	US-08-332-562A-88	Sequence 88, Appl
588	4	19.0	4	90	2	US-08-484-211C-73	Sequence 73, Appl	661	4	19.0	88	2	US-08-562-114B-27	Sequence 27, Appl
589	4	19.0	4	91	3	US-08-906-616-73	Sequence 73, Appl	662	4	19.0	88	3	US-09-053-197A-14	Sequence 14, Appl
590	4	19.0	4	92	3	US-08-817-793-73	Sequence 73, Appl	663	4	19.0	88	4	US-08-729-594A-27	Sequence 27, Appl
591	4	19.0	4	93	4	US-08-485-443B-73	Sequence 73, Appl	664	4	19.0	88	4	US-09-085-761A-14	Sequence 14, Appl
592	4	19.0	4	94	5	US-08-639-075A-73	Sequence 73, Appl	665	4	19.0	89	2	US-08-422-625-2	Sequence 2, Appl
593	4	19.0	4	95	4	US-09-012-431-73	Sequence 73, Appl	666	4	19.0	89	2	US-08-422-333-13	Sequence 13, Appl
594	4	19.0	4	96	4	US-09-012-692-73	Sequence 73, Appl	667	4	19.0	89	4	US-09-203-623-3	Sequence 3, Appl
595	4	19.0	4	97	4	US-08-711-417C-186	Sequence 186, App	668	4	19.0	89	4	US-09-320-574-2	Sequence 2, Appl
596	4	19.0	4	98	4	US-08-711-417C-187	Sequence 187, App	669	4	19.0	89	4	US-08-905-223-433	Sequence 433, App
597	4	19.0	4	99	4	US-08-711-417C-188	Sequence 188, App	670	4	19.0	89	6	5187153-21	Patent No. 5187153
598	4	19.0	4	100	4	US-08-711-417C-189	Sequence 189, App	671	4	19.0	89	6	5223482-23	Patent No. 5223482
599	4	19.0	4	101	4	US-08-906-613-73	Sequence 73, Appl	672	4	19.0	89	6	5466783-3	Patent No. 5466783
600	4	19.0	4	102	4	US-09-187-789-58	Sequence 58, Appl	673	4	19.0	90	6	5220013-24	Patent No. 5220013
601	4	19.0	4	103	5	PCT-US95-1442A-73	Sequence 73, Appl	674	4	19.0	92	4	US-09-334-601-29	Sequence 29, Appl
602	4	19.0	4	104	3	US-08-955-937A-7	Sequence 7, Appl	675	4	19.0	92	4	US-09-334-601-30	Sequence 30, Appl
603	4	19.0	4	105	4	US-09-300-985-7	Sequence 7, Appl	676	4	19.0	93	4	US-09-247-155-114	Sequence 114, App
604	4	19.0	4	106	1	US-08-488-961-3	Sequence 3, Appl	677	4	19.0	94	1	US-08-266-451B-20	Sequence 20, Appl
605	4	19.0	4	107	4	US-08-973-297-3	Sequence 3, Appl	678	4	19.0	94	2	US-08-748-725-20	Sequence 20, Appl
606	4	19.0	4	108	5	PCT-US96-06511-3	Sequence 3, Appl	679	4	19.0	94	3	US-08-488-551B-640	Sequence 640, App
607	4	19.0	4	109	6	US-08-687-702-12	Sequence 12, App	680	4	19.0	96	4	US-08-465-343A-11	Sequence 11, Appl
608	4	19.0	4	110	6	US-08-464-531-115	Sequence 115, App	681	4	19.0	98	1	US-08-118-270-337	Sequence 337, App
609	4	19.0	4	111	2	US-08-461-598-115	Sequence 115, App	682	4	19.0	98	2	US-09-010-928B-20	Sequence 20, Appl
610	4	19.0	4	112	3	US-08-322-137-115	Sequence 115, App	683	4	19.0	98	4	US-09-214-095D-4	Sequence 4, Appl
611	4	19.0	4	113	3			684	4	19.0	98	4		

685	4	19.0	98	5	PCT-US93-08528-337	Sequence 337, App	758	4	19.0	111	3	US-08-545-809A-98	Sequence 98, Appl
686	4	19.0	99	2	US-08-611-510-7	Sequence 7, Appl	759	4	19.0	111	4	US-09-470-614-2	Sequence 2, Appl
687	4	19.0	99	2	US-08-672-345C-91	Sequence 91, Appl	760	4	19.0	112	1	US-08-081-539-72	Sequence 72, Appl
688	4	19.0	99	3	US-08-894-699-34	Sequence 34, Appl	761	4	19.0	112	1	US-08-466-647-72	Sequence 72, Appl
689	4	19.0	99	4	US-09-444-410-34	Sequence 44, Appl	762	4	19.0	112	1	US-09-193-877-4	Sequence 4, Appl
690	4	19.0	99	4	US-09-214-095D-88	Sequence 88, Appl	763	4	19.0	112	4	US-08-857-076-75	Sequence 75, Appl
691	4	19.0	99	4	US-09-214-095D-92	Sequence 92, Appl	764	4	19.0	112	4	US-09-157-370-6	Sequence 6, Appl
692	4	19.0	100	3	US-09-127-424-2	Sequence 2, Appl	765	4	19.0	112	6	5494663-9	Patent No. 5494663
693	4	19.0	100	4	US-09-214-095D-96	Sequence 96, Appl	766	4	19.0	114	1	US-08-436-463-16	Sequence 16, Appl
694	4	19.0	100	4	US-09-495-082-2	Sequence 2, Appl	767	4	19.0	114	2	US-08-637-759B-441	Sequence 441, App
695	4	19.0	102	4	US-09-461-697-441	Sequence 441, App	768	4	19.0	114	3	US-08-871-355A-441	Sequence 441, App
696	4	19.0	103	4	US-09-247-155-111	Sequence 111, App	769	4	19.0	114	4	US-09-203-623-2	Sequence 2, Appl
697	4	19.0	104	4	US-08-936-165A-432	Sequence 432, App	770	4	19.0	114	4	US-09-201-945-441	Sequence 2, Appl
698	4	19.0	105	1	US-07-998-003A-99	Sequence 99, App	771	4	19.0	115	1	US-08-558-735-2	Sequence 12, Appl
699	4	19.0	105	1	US-08-453-274B-99	Sequence 99, App	772	4	19.0	115	2	US-08-481-956A-12	Sequence 25, Appl
700	4	19.0	105	1	US-08-241-853-11	Sequence 11, App	773	4	19.0	115	2	US-08-580-988A-25	Sequence 12, Appl
701	4	19.0	105	1	US-08-326-117B-4	Sequence 4, App	774	4	19.0	115	2	US-08-629-291A-12	Sequence 12, Appl
702	4	19.0	105	1	US-08-453-695A-99	Sequence 99, App	775	4	19.0	115	2	US-08-658-358B-12	Sequence 2, Appl
703	4	19.0	105	1	US-08-268-161A-99	Sequence 99, App	776	4	19.0	115	4	US-08-906-480-2	Sequence 2, Appl
704	4	19.0	105	2	US-08-850-917-11	Sequence 11, App	777	4	19.0	115	6	5215889-3	Patent No. 5215889
705	4	19.0	105	2	US-08-301-915-2	Sequence 2, App	778	4	19.0	116	1	US-08-478-039-67	Sequence 67, App
706	4	19.0	105	2	US-08-464-517-28	Sequence 28, App	779	4	19.0	116	1	US-08-478-039-74	Sequence 74, App
707	4	19.0	105	2	US-08-453-702A-99	Sequence 99, App	780	4	19.0	116	1	US-08-478-039-103	Sequence 103, App
708	4	19.0	105	2	US-08-246-361A-28	Sequence 28, App	781	4	19.0	116	1	US-08-476-349A-67	Sequence 67, App
709	4	19.0	105	3	US-08-982-139-4	Sequence 4, App	782	4	19.0	116	1	US-08-476-349A-74	Sequence 74, App
710	4	19.0	105	3	US-08-488-551B-639	Sequence 639, App	783	4	19.0	116	1	US-08-476-349A-103	Sequence 103, App
711	4	19.0	105	3	US-08-424-694A-2	Sequence 2, App	784	4	19.0	116	1	US-09-228-986-96	Sequence 96, App
712	4	19.0	105	3	US-08-463-772-28	Sequence 28, App	785	4	19.0	117	4	US-08-969-683A-62	Sequence 62, App
713	4	19.0	105	3	US-08-867-381A-4	Sequence 4, App	786	4	19.0	117	4	5514582-15	Patent No. 5514582
714	4	19.0	105	4	US-09-099-639-99	Sequence 99, App	787	4	19.0	118	4	US-09-370-253-18	Sequence 18, App
715	4	19.0	105	4	US-09-521-144-4	Sequence 4, App	788	4	19.0	119	2	US-08-737-560A-10	Sequence 10, App
716	4	19.0	105	5	PCT-US93-05000-28	Sequence 28, App	789	4	19.0	119	3	US-09-023-082A-80	Sequence 80, App
717	4	19.0	105	5	PCT-US93-12589-99	Sequence 99, App	790	4	19.0	119	3	US-08-545-809A-93	Sequence 93, App
718	4	19.0	105	5	PCT-US95-08071-99	Sequence 99, App	791	4	19.0	119	4	US-08-545-809A-93	Sequence 93, App
719	4	19.0	106	2	US-08-440-354-2	Sequence 2, App	792	4	19.0	120	1	US-08-858-207A-397	Sequence 397, App
720	4	19.0	106	2	US-08-463-087-2	Sequence 2, App	793	4	19.0	120	1	US-08-253-155A-37	Sequence 37, App
721	4	19.0	106	2	US-08-820-825-14	Sequence 14, App	794	4	19.0	120	2	US-08-290-592E-31	Sequence 31, App
722	4	19.0	106	2	US-08-820-825-14	Sequence 14, App	795	4	19.0	120	2	US-08-290-592E-32	Sequence 32, App
723	4	19.0	106	2	US-09-307-817-14	Sequence 14, App	796	4	19.0	120	2	US-08-553-501A-90	Sequence 90, App
724	4	19.0	106	4	US-09-452-817-2	Sequence 2, App	797	4	19.0	120	2	US-08-553-501A-91	Sequence 91, App
725	4	19.0	106	6	5204252-6	Patent No. 5204252	798	4	19.0	120	2	US-08-647-144-2	Sequence 2, App
726	4	19.0	106	6	5204252-7	Patent No. 5204252	799	4	19.0	120	3	US-09-203-231-90	Sequence 90, App
727	4	19.0	107	1	US-08-409-731A-2	Sequence 2, App	800	4	19.0	120	3	US-09-203-231-91	Sequence 91, App
728	4	19.0	107	2	US-08-470-298B-2	Sequence 2, App	801	4	19.0	120	4	US-09-025-769B-23	Sequence 23, App
729	4	19.0	107	2	US-09-023-073A-2	Sequence 2, App	802	4	19.0	120	4	US-09-199-637A-219	Sequence 219, App
730	4	19.0	107	3	US-08-388-353-640	Sequence 2, App	803	4	19.0	120	5	PCT-US95-10053-28	Sequence 28, App
731	4	19.0	107	4	US-09-361-737-2	Sequence 640, App	804	4	19.0	120	5	PCT-US95-10053-29	Sequence 29, App
732	4	19.0	107	4	US-08-936-165A-430	Sequence 2, App	805	4	19.0	120	5	PCT-US96-09448-31	Sequence 31, App
733	4	19.0	108	1	US-08-436-463-15	Sequence 430, App	806	4	19.0	120	5	PCT-US96-09448-32	Sequence 32, App
734	4	19.0	108	1	US-08-436-463-15	Sequence 15, App	807	4	19.0	121	2	US-08-483-636-64	Sequence 64, App
735	4	19.0	108	3	US-08-388-353-639	Sequence 17, App	808	4	19.0	121	4	US-09-025-769B-37	Sequence 37, App
736	4	19.0	108	4	US-09-240-274-51	Sequence 51, App	809	4	19.0	121	4	US-09-025-769B-61	Sequence 61, App
737	4	19.0	108	4	US-09-240-274-52	Sequence 52, App	810	4	19.0	122	1	US-07-734-534A-1	Sequence 1, App
738	4	19.0	108	4	US-09-240-274-53	Sequence 53, App	811	4	19.0	122	1	US-08-276-852-143	Sequence 143, App
739	4	19.0	108	4	US-09-240-274-166	Sequence 166, App	812	4	19.0	122	1	US-08-487-748A-17	Sequence 17, App
740	4	19.0	109	1	US-08-436-463-8	Sequence 8, App	813	4	19.0	122	1	US-08-436-463-14	Sequence 14, App
741	4	19.0	109	1	US-08-028-253-8	Sequence 8, App	814	4	19.0	122	1	US-08-899-575-143	Sequence 143, App
742	4	19.0	109	2	US-08-672-345C-1	Sequence 2, App	815	4	19.0	122	1	US-08-716-588-3	Sequence 3, App
743	4	19.0	109	2	US-08-672-345C-3	Sequence 3, App	816	4	19.0	122	1	US-08-899-575-143	Sequence 3, App
744	4	19.0	109	2	US-08-672-345C-4	Sequence 4, App	817	4	19.0	122	2	US-08-706-741B-51	Sequence 51, App
745	4	19.0	109	2	US-09-121-179-1	Sequence 1, App	818	4	19.0	122	2	US-08-966-444-3	Sequence 51, App
746	4	19.0	109	2	US-09-214-095D-1	Sequence 1, App	819	4	19.0	122	3	US-08-924-695A-51	Sequence 17, App
747	4	19.0	109	4	US-09-214-095D-2	Sequence 2, App	820	4	19.0	122	3	US-08-398-633-17	Sequence 17, App
748	4	19.0	109	4	US-09-214-095D-3	Sequence 3, App	821	4	19.0	122	3	US-08-480-070C-17	Sequence 17, App
749	4	19.0	109	4	US-09-214-095D-4	Sequence 4, App	822	4	19.0	122	3	US-08-829-525-17	Sequence 17, App
750	4	19.0	109	4	US-09-214-095D-121	Sequence 121, App	823	4	19.0	122	4	US-08-609-583A-17	Sequence 17, App
751	4	19.0	110	1	US-08-122-546-14	Sequence 14, App	824	4	19.0	122	4	US-08-937-399-17	Sequence 17, App
752	4	19.0	110	1	US-08-422-613-2	Sequence 2, App	825	4	19.0	122	5	PCT-US95-08763-143	Sequence 143, App
753	4	19.0	110	3	US-08-764-938-14	Sequence 14, App	826	4	19.0	123	3	US-08-483-577A-159	Sequence 159, App
754	4	19.0	110	3	US-09-131-052-14	Sequence 14, App	827	4	19.0	123	4	US-08-897-438-159	Sequence 159, App
755	4	19.0	110	4	US-09-131-053A-14	Sequence 14, App	828	4	19.0	124	2	US-08-743-200-12	Sequence 12, App
756	4	19.0	110	4	US-08-945-997C-20	Sequence 20, App	829	4	19.0	124	3	US-08-991-890-5	Sequence 5, App
757	4	19.0	110	4	US-08-905-223-336	Sequence 336, App	830	4	19.0	124	4	US-08-751-359-14	Sequence 14, App

831	4	19.0	134	4	US-08-907-145-14	Sequence 14, Appl	904	4	19.0	139	4	US-08-857-076-13	Sequence 13, Appl
832	4	19.0	124	4	US-08-858-207A-498	Sequence 498, App	905	4	19.0	139	4	US-08-874-102-36	Sequence 36, Appl
833	4	19.0	125	1	US-08-478-039-69	Sequence 69, Appl	906	4	19.0	139	4	US-08-936-165A-409	Sequence 409, App
834	4	19.0	126	1	US-08-476-349A-69	Sequence 69, Appl	907	4	19.0	140	2	US-08-483-636-4	Sequence 4, Appl
835	4	19.0	125	1	US-08-276-852-142	Sequence 142, App	908	4	19.0	140	2	US-08-483-632-4	Sequence 4, Appl
836	4	19.0	126	1	US-08-899-575-142	Sequence 142, App	909	4	19.0	141	2	US-08-345-321-10	Sequence 10, Appl
837	4	19.0	126	1	US-08-899-575-142	Sequence 142, App	910	4	19.0	141	2	US-08-483-636-10	Sequence 10, Appl
838	4	19.0	126	1	US-08-751-359-11	Sequence 11, Appl	911	4	19.0	141	2	US-08-483-636-12	Sequence 12, Appl
839	4	19.0	126	4	US-08-907-146-11	Sequence 11, Appl	912	4	19.0	141	2	US-08-483-632-10	Sequence 10, Appl
840	4	19.0	126	5	PCT-US95-08743-142	Sequence 142, App	913	4	19.0	141	2	US-08-483-632-12	Sequence 12, Appl
841	4	19.0	126	2	PCT-US96-03916-15	Sequence 15, Appl	914	4	19.0	142	4	US-08-456-830-31	Sequence 31, Appl
842	4	19.0	127	2	US-08-637-759B-334	Sequence 334, App	915	4	19.0	142	4	US-08-456-830-47	Sequence 47, Appl
843	4	19.0	127	3	US-08-871-355A-334	Sequence 334, App	916	4	19.0	142	4	US-08-456-830-59	Sequence 59, Appl
844	4	19.0	127	4	US-09-201-945-334	Sequence 334, App	917	4	19.0	142	6	5242807-2	Patent No. 5242807
845	4	19.0	128	2	US-08-820-170A-7	Sequence 7, Appl	918	4	19.0	143	1	US-08-436-463-2	Sequence 2, Appl
846	4	19.0	128	2	US-08-318-837-51	Sequence 51, Appl	919	4	19.0	143	3	US-09-129-075-12	Sequence 12, Appl
847	4	19.0	128	3	US-09-055-699-7	Sequence 7, Appl	920	4	19.0	143	4	US-08-868-699A-4	Sequence 4, Appl
848	4	19.0	128	4	US-09-273-565-7	Sequence 7, Appl	921	4	19.0	143	4	US-09-757-014-4	Sequence 4, Appl
849	4	19.0	128	4	US-08-348-548-2	Sequence 2, Appl	922	4	19.0	144	2	US-08-313-185-60	Sequence 60, Appl
850	4	19.0	128	4	US-08-885-469-2	Sequence 2, Appl	923	4	19.0	144	3	US-08-082-614A-60	Sequence 60, Appl
851	4	19.0	128	4	US-09-565-538-7	Sequence 7, Appl	924	4	19.0	144	4	US-09-045-373-9	Sequence 9, Appl
852	4	19.0	128	4	US-09-228-986-102	Sequence 102, App	925	4	19.0	144	4	US-09-319-989-4	Sequence 4, Appl
853	4	19.0	128	5	PCT-US95-15716-2	Sequence 2, Appl	926	4	19.0	145	2	US-08-866-599A-20	Sequence 20, Appl
854	4	19.0	129	3	US-09-187-049-5	Sequence 5, Appl	927	4	19.0	147	1	US-08-276-151-4	Sequence 4, Appl
855	4	19.0	129	4	US-09-476-482-8	Sequence 8, Appl	928	4	19.0	147	2	US-08-385-590A-4	Sequence 4, Appl
856	4	19.0	129	6	5428135-6	Patent No. 5428135	929	4	19.0	147	3	US-09-021-520-4	Sequence 4, Appl
857	4	19.0	130	1	US-08-478-039-66	Sequence 66, App	930	4	19.0	147	4	US-09-388-917-5	Sequence 5, Appl
858	4	19.0	130	1	US-08-478-039-106	Sequence 106, App	931	4	19.0	148	1	US-08-233-788A-43	Sequence 43, Appl
859	4	19.0	130	1	US-08-476-349A-66	Sequence 66, App	932	4	19.0	148	2	US-08-322-625-18	Sequence 18, Appl
860	4	19.0	130	1	US-08-476-349A-106	Sequence 106, App	933	4	19.0	148	2	US-08-466-961A-18	Sequence 18, Appl
861	4	19.0	130	2	US-08-630-822A-90	Sequence 90, Appl	934	4	19.0	148	2	US-08-645-193B-13	Sequence 13, Appl
862	4	19.0	130	2	US-09-005-069-90	Sequence 90, Appl	935	4	19.0	149	3	US-08-808-599A-40	Sequence 40, Appl
863	4	19.0	131	2	US-08-477-451-30	Sequence 30, Appl	936	4	19.0	150	6	5310729-2	Patent No. 5310729
864	4	19.0	131	2	US-08-820-825-15	Sequence 15, Appl	937	4	19.0	151	1	US-08-485-455D-51	Sequence 51, Appl
865	4	19.0	131	2	US-08-850-910A-39	Sequence 39, Appl	938	4	19.0	151	2	US-08-722-050-5	Sequence 5, Appl
866	4	19.0	131	2	US-08-850-910A-41	Sequence 41, Appl	939	4	19.0	151	2	US-08-482-130C-51	Sequence 51, Appl
867	4	19.0	131	2	US-08-850-910A-43	Sequence 43, Appl	940	4	19.0	151	2	US-08-484-211C-51	Sequence 51, Appl
868	4	19.0	131	2	US-08-850-910A-46	Sequence 46, Appl	941	4	19.0	151	3	US-08-906-769-51	Sequence 51, Appl
869	4	19.0	131	4	US-09-307-817-15	Sequence 15, Appl	942	4	19.0	151	3	US-08-906-616-51	Sequence 51, Appl
870	4	19.0	132	2	US-08-470-298B-12	Sequence 12, Appl	943	4	19.0	151	4	US-08-817-795-51	Sequence 51, Appl
871	4	19.0	132	2	US-08-850-910A-47	Sequence 47, Appl	944	4	19.0	151	4	US-08-485-443B-51	Sequence 51, Appl
872	4	19.0	132	4	US-09-159-106-15	Sequence 15, Appl	945	4	19.0	151	4	US-08-639-075A-51	Sequence 51, Appl
873	4	19.0	132	4	US-09-043-646-2	Sequence 2, Appl	946	4	19.0	151	4	US-09-012-431-51	Sequence 51, Appl
874	4	19.0	134	1	US-08-328-152A-13	Sequence 13, Appl	947	4	19.0	151	4	US-09-012-692-51	Sequence 51, Appl
875	4	19.0	135	1	US-08-409-731A-8	Sequence 8, Appl	948	4	19.0	151	4	US-08-906-613-51	Sequence 51, Appl
876	4	19.0	135	1	US-08-336-618-20	Sequence 20, Appl	949	4	19.0	151	5	PCT-US95-14442A-51	Sequence 51, Appl
877	4	19.0	135	2	US-08-470-298B-8	Sequence 8, Appl	950	4	19.0	154	4	US-09-193-104-26	Sequence 26, Appl
878	4	19.0	135	2	US-08-820-825-2	Sequence 2, Appl	951	4	19.0	154	4	US-09-247-155-110	Sequence 110, App
879	4	19.0	135	2	US-08-847-724-4	Sequence 4, Appl	952	4	19.0	154	6	5183734-1	Patent No. 5183734
880	4	19.0	135	2	US-08-847-724-6	Sequence 6, Appl	953	4	19.0	156	3	US-08-826-246-10	Sequence 10, Appl
881	4	19.0	135	2	US-09-023-073A-8	Sequence 8, Appl	954	4	19.0	156	3	US-08-944-495-10	Sequence 10, Appl
882	4	19.0	135	3	US-08-899-031-1	Sequence 1, Appl	955	4	19.0	156	3	US-09-126-640-11	Sequence 11, Appl
883	4	19.0	135	3	US-08-899-031-3	Sequence 3, Appl	956	4	19.0	156	4	US-08-925-588-10	Sequence 10, Appl
884	4	19.0	135	3	US-08-899-031-4	Sequence 4, Appl	957	4	19.0	156	4	US-09-056-556-227	Sequence 227, App
885	4	19.0	135	4	US-09-091-117-1	Sequence 1, Appl	958	4	19.0	156	4	US-09-288-282A-11	Sequence 11, Appl
886	4	19.0	135	4	US-09-307-817-2	Sequence 2, Appl	959	4	19.0	157	1	US-08-257-399-4	Sequence 4, Appl
887	4	19.0	135	4	US-09-361-737-8	Sequence 8, Appl	960	4	19.0	157	2	US-08-811-949-41	Sequence 41, Appl
888	4	19.0	136	1	US-08-037-579A-7	Sequence 7, Appl	961	4	19.0	158	1	US-08-992-176-5	Sequence 5, Appl
889	4	19.0	136	2	US-08-477-451-31	Sequence 31, Appl	962	4	19.0	158	1	US-07-778-156-3	Sequence 3, Appl
890	4	19.0	136	3	US-08-601-184-7	Sequence 7, Appl	963	4	19.0	158	2	US-08-422-166-3	Sequence 3, Appl
891	4	19.0	136	4	US-09-136-879-5	Sequence 5, Appl	964	4	19.0	158	3	US-09-053-197A-19	Sequence 19, Appl
892	4	19.0	137	3	US-09-053-197A-31	Sequence 31, Appl	965	4	19.0	158	4	US-09-215-221-22	Sequence 22, Appl
893	4	19.0	137	4	US-09-456-830-41	Sequence 41, Appl	966	4	19.0	158	4	US-09-267-177-10	Sequence 10, Appl
894	4	19.0	137	4	US-09-085-761A-31	Sequence 31, Appl	967	4	19.0	158	4	US-09-085-761A-19	Sequence 19, Appl
895	4	19.0	138	1	US-08-686-878A-14	Sequence 14, Appl	968	4	19.0	159	1	US-08-414-926A-2	Sequence 2, Appl
896	4	19.0	138	3	US-09-191-647-4	Sequence 4, Appl	969	4	19.0	159	1	US-08-414-926A-21	Sequence 21, Appl
897	4	19.0	138	4	US-09-540-245A-4	Sequence 4, Appl	970	4	19.0	159	2	US-08-926-922-2	Sequence 2, Appl
898	4	19.0	138	4	US-09-540-153-4	Sequence 4, Appl	971	4	19.0	159	2	US-08-926-922-21	Sequence 21, Appl
899	4	19.0	139	2	US-08-553-501A-59	Sequence 59, Appl	972	4	19.0	159	3	US-09-253-682-2	Sequence 2, Appl
900	4	19.0	139	3	US-08-950-720A-15	Sequence 15, Appl	973	4	19.0	159	3	US-09-253-682-21	Sequence 21, Appl
901	4	19.0	139	3	US-08-991-890-2	Sequence 2, Appl	974	4	19.0	159	3	US-08-991-890-4	Sequence 4, Appl
902	4	19.0	139	3	US-09-205-231-59	Sequence 59, Appl	975	4	19.0	159	4	US-09-527-657-2	Sequence 2, Appl
903	4	19.0	139	4	US-08-444-818-172	Sequence 172, App	976	4	19.0	159	4	US-09-527-657-21	Sequence 21, Appl

977 4 19.0 160 1 US-08-014-153D-8
978 4 19.0 160 2 US-08-883-070-7
979 4 19.0 161 2 US-08-997-080-44
980 4 19.0 161 2 US-08-997-080-70
981 4 19.0 161 2 US-08-997-362-44
982 4 19.0 161 2 US-08-997-362-70
983 4 19.0 161 3 US-08-873-970-44
984 4 19.0 161 3 US-08-873-970-70
985 4 19.0 161 4 US-09-095-855-44
986 4 19.0 161 4 US-09-095-855-70
987 4 19.0 161 4 US-08-703-347A-44
988 4 19.0 161 4 US-09-324-542-44
989 4 19.0 161 4 US-09-324-542-70
990 4 19.0 162 1 US-08-362-453-15
991 4 19.0 162 4 US-08-800-682-7
992 4 19.0 163 1 US-08-469-667-11
993 4 19.0 163 4 US-09-224-110-11
994 4 19.0 163 5 PCT-US95-07289-11
995 4 19.0 164 1 US-08-357-125-4
996 4 19.0 164 1 US-08-318-193-77
997 4 19.0 165 1 US-08-024-330-1
998 4 19.0 165 1 US-07-952-840-1
999 4 19.0 165 1 US-08-145-995A-11
1000 4 19.0 165 5 PCT-US95-03866-5

ALIGNMENTS

RESULT 1
US-08-482-918-56
; Sequence 56, Application US/08482918
; Patent No. 6207417
; GENERAL INFORMATION:
; APPLICANT: Zsebo, Krisztina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Martin, Francis H.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,918
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 01017/33005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-482-918-56

Query Match 28.6%; Score 6; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 34; 0; Indels 0; Gaps 0;
Matches 6; Conservative 0; Mismatches 0;
QY 7 FFRLFN 12
Db 139 FFRLFN 144
|||||
RESULT 2
US-09-224-681-56
; Sequence 56, Application US/09224681
; Patent No. 6207454
; GENERAL INFORMATION:
; APPLICANT: Zsebo, Krisztina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; APPLICANT: Martin, Francis H.
; TITLE OF INVENTION: Method for Enhancing the Efficiency of Gene
; TRANSFER WITH STEM CELL FACTOR (SCF) POLYPEPTIDE
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/224,681
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/005,893
; FILING DATE: 12-JAN-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/449,653
; FILING DATE: 24-MAY-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/982,255
; FILING DATE: 25-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/589,701
; FILING DATE: 01-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/573,616
; FILING DATE: 24-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/537,198
; FILING DATE: 11-JUN-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/422,383
; FILING DATE: 16-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.
; REGISTRATION NUMBER: 36,107
; REFERENCE/DOCKET NUMBER: 01017/35199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-224-681-56

Query Match 28.6%; Score 6; DB 4; Length 282;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEN 12
|11111|
Db 139 FFRLEN 144

RESULT 3

US-08-336-728A-56
; Sequence 56, Application US/08336728A
; Patent No. 6207802
; GENERAL INFORMATION:
; APPLICANT: Zsebo, Krisztina M.
; APPLICANT: Bosselman, Robert A.
; APPLICANT: Suggs, Sidney V.
; TITLE OF INVENTION: Stem Cell Factor
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,728A
; FILING DATE: 09-NOV-1994
; CLASSIFICATION: 424

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/982,255
; FILING DATE: 25-NOV-1992

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/589,701
; FILING DATE: 01-OCT-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/573,616
; FILING DATE: 24-AUG-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/537,198
; FILING DATE: 11-JUN-1990

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/422,383
; FILING DATE: 16-OCT-1989

ATTORNEY/AGENT INFORMATION:
; NAME: Clough, David W.

REGISTRATION NUMBER: 36,107
REFERENCE/DOCKET NUMBER: 01017/32956

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:
; LENGTH: 282 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-336-728A-56

Query Match

28.6%; Score 6; DB 4; Length 282;

Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLEN 12
|11111|
Db 139 FFRLEN 144

RESULT 4

US-09-268-992-51
; Sequence 51, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 51
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-992-51

Query Match 28.6%; Score 6; DB 4;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
|11111|
Db 174 LFNRSF 179

RESULT 5

US-09-268-992-64
; Sequence 64, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 64
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-992-64

Query Match 28.6%; Score 6; DB 4; Length 466;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
| | | | |
Db 194 LFNRSF 199

RESULT 6

US-09-268-992-4
; Sequence 4, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-992-4

Query Match 28.6%; Score 6; DB 4; Length 477;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
| | | | |
Db 225 LFNRSF 230

RESULT 7

US-09-268-992-2
; Sequence 2, Application US/09268992
; Patent No. 6342351
; GENERAL INFORMATION:
; APPLICANT: Chen, H.
; APPLICANT: Freimer, N.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; AND TREATING CHROMOSOME-18p RELATED DISORDERS
; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/268,992
; CURRENT FILING DATE: 1999-03-16
; EARLIER APPLICATION NUMBER: 09/236,134
; EARLIER FILING DATE: 1999-01-22
; EARLIER APPLICATION NUMBER: 60/106,056
; EARLIER FILING DATE: 1998-10-28
; EARLIER APPLICATION NUMBER: 60/088,312
; EARLIER FILING DATE: 1998-06-05
; EARLIER APPLICATION NUMBER: 60/078,044
; EARLIER FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-268-992-2

Query Match 28.6%; Score 6; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
| | | | |
Db 223 LFNRSF 228

RESULT 8

US-07-952-853-6
; Sequence 6, Application US/07952853
; Patent No. 5863783
; GENERAL INFORMATION:
; APPLICANT: Van Heuvel, Margaretha
; APPLICANT: Bakhuis, Janna G.
; APPLICANT: Coutel, Yves
; APPLICANT: Harder, Abraham
; APPLICANT: De Graaff, Leendert H.
; APPLICANT: Flippin, Michel J. A.
; APPLICANT: Van Der Veen, Peter
; APPLICANT: Visser, Jacob
; APPLICANT: Andreoli, Peter M.
; TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
; MOLECULES
; TITLE OF INVENTION: ENCODING ARABINAN-DEGRADING ENZYMES OF FUNGAL
; ORIGIN
; TITLE OF INVENTION: ORIGIN
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/952,853
; FILING DATE: 19921125
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 246152003500
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-813-5600
; TELEFAX: 415-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 499 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-952-853-6

Query Match 28.6%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLST 6
| | | | |
Db 264 SGSLST 269

RESULT 9

US-08-914-848-6
; Sequence 6, Application US/08914848
; Patent No. 5989887

GENERAL INFORMATION:
APPLICANT: Van Heuvel, Margaretha
APPLICANT: Bakhuis, Janna G.
APPLICANT: Cotel, Yves
APPLICANT: Harder, Abraham
APPLICANT: De Graaff, Leendert H.
APPLICANT: Flippin, Michel J. A.
APPLICANT: Van Der Veen, Peter
APPLICANT: Visser, Jacob
APPLICANT: Andreoli, Peter M.
TITLE OF INVENTION: CLONING AND EXPRESSION OF DNA
TITLE OF INVENTION: MOLECULES ENCODING ARABINAN-DEGRADING ENZYMES OF FUNGAL
TITLE OF INVENTION: ORIGIN
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,848
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/952,853
FILING DATE: 25-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 246152003500
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-914-848-6

Query Match 28.6%; Score 6; DB 2; Length 499;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SCSLST 6
Db 264 SCSLST 269

RESULT 10
US-09-014-969-11
Sequence 11, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallie, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-014-969-11

Query Match 28.6%; Score 6; DB 2; Length 635;
Best Local Similarity 100.0%; Pred. No. 70;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SCSLST 6
Db 361 SCSLST 366

RESULT 11
US-07-745-206A-7
Sequence 7, Application US/07745206A
Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1968 amino acids
TYPE: AMINO ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-07-745-206A-7

Query Match 28.6%; Score 6; DB 1; Length 1968;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
|||||

Db 1351 TFFRLF 1356

RESULT 12

US-08-455-543A-45
Sequence 45, Application US/08455543A
Patent No. 5792846

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
FILING DATE: May 31, 1995
PRIOR APPLICATION NUMBER: US/08/455,543A
FILING DATE: April 4, 1994
APPLICATION NUMBER: 08/223,305
FILING DATE: April 4, 1994
PRIOR APPLICATION DATA:
FILING DATE: April 10, 1992
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
FILING DATE: April 10, 1992
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
FILING DATE: 30-NOV-1990
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
FILING DATE: 20-FEB-1990
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
FILING DATE: 04-APR-1989
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
FILING DATE: 04-APR-1989
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
FILING DATE: 04-APR-1989
APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-52517
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1968 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-455-543A-45

Query Match 28.6%; Score 6; DB 1; Length 1968;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
|||||

Db 1351 TFFRLF 1356

RESULT 13

US-08-223-305C-45
Sequence 45, Application US/08223305C
Patent No. 5851824

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
FILING DATE: April 4, 1994
APPLICATION NUMBER: US/08/223,305C
FILING DATE: April 10, 1992
APPLICATION NUMBER: 07/868,354
FILING DATE: April 10, 1992
PRIOR APPLICATION DATA:
FILING DATE: 15-AUG-1991
APPLICATION NUMBER: US 07/745,206
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
FILING DATE: 30-NOV-1990
APPLICATION NUMBER: US 07/620,250
FILING DATE: 30-NOV-1990
PRIOR APPLICATION DATA:
FILING DATE: 20-FEB-1990
APPLICATION NUMBER: US 07/482,384
FILING DATE: 20-FEB-1990
PRIOR APPLICATION DATA:
FILING DATE: 04-APR-1989
APPLICATION NUMBER: US 07/603,751
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
FILING DATE: 04-APR-1989
APPLICATION NUMBER: WO PCT/US89/01408
FILING DATE: 04-APR-1989
PRIOR APPLICATION DATA:
ATTORNEY/AGENT INFORMATION:

APPLICATION NUMBER: US 07/176,899
FILING DATE: 04-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 52516 (P519739)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 45:
SEQUENCE CHARACTERISTICS:
LENGTH: 1968 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-223-305C-45

Query Match 28.6%; Score 6; DB 2; Length 1968;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
|||||
Db 1351 TFFRLF 1356

RESULT 14
US-08-311-363-7
Sequence 7, Application US/08311363
Patent No. 5876958
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert
TITLE OF INVENTION: Human Calcium Channel Compositions and
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Brown, Martin, Haller & McClain
STREET: 1660 Union Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101-2926
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/311,363
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/745,206
FILING DATE: 15-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Seidman, Stephanie L.
REGISTRATION NUMBER: 33,779
REFERENCE/DOCKET NUMBER: 6362-51506
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619)238-0999
TELEFAX: (619)238-0062
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1968 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-311-363-7
Query Match 28.6%; Score 6; DB 2; Length 1968;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 TFFRLF 11
|||||
Db 1351 TFFRLF 1356
RESULT 15
US-07-745-206A-2
Sequence 2, Application US/07745206A
Patent No. 5429921
GENERAL INFORMATION:
APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESS: Fitch, Even, Tabin & Flannery
STREET: 135 S. LaSalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2161 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-745-206A-2

Query Match 28.6%; Score 6; DB 1; Length 2161;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
|||||
Db 1316 TFFRLF 1321

Search completed: November 5, 2002, 11:07:06
Job time: 16.8955 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 11:02:02 ; Search time 8.14925 Seconds
(without alignments)
247.615 Million cell updates/sec

Title: US-09-833-017b-4

Perfect score: 21
Sequence: 1 SGSLSTFFRLFNRSFTQALGK 21

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 283138 seqs, 96089334 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_71.*

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	7	33.3	230	2 S46737	hypothetical prote
2	7	33.3	253	2 S70367	stem cell factor 5
3	7	33.3	287	2 S70366	stem cell factor 1
4	7	33.3	472	2 AD2284	hypothetical prote
5	7	33.3	565	2 A10479	probable membrane
6	6	28.6	43	2 A97416	hypothetical prote
7	6	28.6	69	2 T18094	hypothetical prote
8	6	28.6	122	2 E75387	hypothetical prote
9	6	28.6	128	2 A11330	hypothetical prote
10	6	28.6	134	2 T17655	hypothetical prote
11	6	28.6	155	2 T70370	hypothetical prote
12	6	28.6	181	2 S78372	EST3 protein - yea
13	6	28.6	186	2 A56956	GTP-binding protel
14	6	28.6	197	2 A44363	voltage-gated dihy
15	6	28.6	202	2 H81731	CDP-diacylglycerol
16	6	28.6	213	2 E71511	hypothetical prote
17	6	28.6	214	2 G70348	conserved hypothet
18	6	28.6	271	2 G84030	hema concentration
19	6	28.6	283	2 C23660	calcium channel pr
20	6	28.6	285	2 E83795	ABC transporter (A
21	6	28.6	287	2 JN0637	stem cell factor p
22	6	28.6	287	2 D97241	phosphohydrolase f
23	6	28.6	288	2 D82578	conserved hypothet
24	6	28.6	289	2 T28922	hypothetical prote
25	6	28.6	294	2 A23660	calcium channel pr
26	6	28.6	294	2 B23660	calcium channel pr
27	6	28.6	295	2 S05006	calcium channel al
28	6	28.6	295	2 F98205	probable transcrip
29	6	28.6	295	2 AC3081	transcription requ

enhancer-binding p
transcription fact
probable transcrip
conserved site-spec
hypothetical prote
hypothetical prote
hypothetical prote
transcription fact
Yop translocation
type III secretion
hypothetical prote
hypothetical prote
probable glycosyl
probable colanic a
probable protein l
hypothetical prote
probable glycosyl
probable glycosyl
alpha-L-arabinofur
hypothetical prote
hypothetical 62.1
probable 2-compone
probable 2-compone
probable two-compo
protein B0212.2 [i
probable cytochrom
protein R119.7 [im
hypothetical prote
hypothetical prote
DNA-directed RNA p
methyilmalonyl-CoA
hypothetical prote
hypothetical prote
ABC transporter, H
PMT6 protein - yea
probable dna repl
probable enzyme [i
yagx protein - Esc
probable enzyme ya
hypothetical prote
myosin-like protei
hypothetical prote
DNA-directed RNA p
hypothetical prote
hypothetical prote
calcium channel al
calcium channel al
voltage-dependent
voltage-dependent
integrin beta-4 ch
exonuclease ABC c
integrin-beta-4 ch
calcium channel pr
serine/threonine k
l7123.15 protein -
voltage-dependent
voltage-dependent
calcium channel al
calcium channel pr
calcium channel al
voltage-dependent
calcium channel pr
hypothetical prote
toxin-like outer m
CPY protein - midy
hypothetical prote
myosin heavy chain
cathepsin D (EC 3
actin alpha, smoot
hypothetical prote
hypothetical prote

103	5	23.8	53	2	F82662	hypothetical prote	176	5	23.8	131	2	S67869	probable membrane
104	5	23.8	53	2	AB1896	hypothetical prote	177	5	23.8	133	2	D69836	hypothetical prote
105	5	23.8	54	2	I46468	actin - sheep (fra	178	5	23.8	134	2	I43561	probable transpos
106	5	23.8	66	2	T29601	hypothetical prote	179	5	23.8	134	2	S72727	dedA protein - Myc
107	5	23.8	67	2	I46473	alpha-actin - rabb	180	5	23.8	135	2	S32386	T-cell receptor be
108	5	23.8	67	2	E81042	hypothetical prote	181	5	23.8	135	2	A36828	or10 protein - Au
109	5	23.8	67	2	E69447	hypothetical prote	182	5	23.8	136	2	AG2009	transcription regu
110	5	23.8	67	2	T07256	hypothetical prote	183	5	23.8	137	2	A28258	actin 5C - fruit f
111	5	23.8	76	2	B75271	hypothetical prote	184	5	23.8	137	2	A82544	transcription regu
112	5	23.8	77	2	T14595	hypothetical prote	185	5	23.8	139	2	A71190	hypothetical prote
113	5	23.8	78	2	S17872	translation initia	186	5	23.8	140	2	S73680	ME337 homolog P02
114	5	23.8	79	2	H82705	hypothetical prote	187	5	23.8	140	2	H94405	hypothetical prote
115	5	23.8	80	2	E711430	hypothetical prote	188	5	23.8	141	2	S30832	hypothetical prote
116	5	23.8	82	2	T18102	hypothetical prote	189	5	23.8	143	2	JC7309	agglutinin I - Jer
117	5	23.8	83	4	S53038	hypothetical prote	190	5	23.8	146	2	T06471	core protein - gar
118	5	23.8	84	2	E90475	hypothetical prote	191	5	23.8	147	2	A81454	probable integral
119	5	23.8	85	2	T26333	hypothetical prote	192	5	23.8	147	2	F83801	chorismate mutase
120	5	23.8	87	1	A36437	flavocytochrome c,	193	5	23.8	148	2	E64939	hypothetical prote
121	5	23.8	89	2	E83141	hypothetical prote	194	5	23.8	148	2	B90941	hypothetical prote
122	5	23.8	90	2	A54937	brain type III sod	195	5	23.8	148	2	D87687	conserved hypotet
123	5	23.8	92	2	B45837	Ig gamma-1 chain C	196	5	23.8	148	2	D87687	hypothetical membr
124	5	23.8	92	2	S57260	nitrite reductase	197	5	23.8	149	2	C90824	Spec2c protein - s
125	5	23.8	94	2	H42518	Al5L 11K protein -	198	5	23.8	151	1	MCUR2C	hypothetical prote
126	5	23.8	94	2	G36849	Al5L protein - var	199	5	23.8	151	2	D87548	hypothetical prote
127	5	23.8	94	2	T28557	hypothetical prote	200	5	23.8	154	2	H82810	bacterioferritin X
128	5	23.8	94	2	E72165	Al6L protein - var	201	5	23.8	154	2	H75031	hypothetical prote
129	5	23.8	95	2	C69449	conserved hypotet	202	5	23.8	154	2	D83516	hypothetical prote
130	5	23.8	97	2	S26652	Ig gamma-1 chain C	203	5	23.8	155	2	AG0945	actin homolog FAT,
131	5	23.8	97	2	E95064	conserved hypotet	204	5	23.8	155	2	F71004	hypothetical prote
132	5	23.8	97	2	G64454	hypothetical prote	205	5	23.8	155	2	A71693	hypothetical prote
133	5	23.8	97	2	G97931	conserved hypotet	206	5	23.8	156	2	T16531	hypothetical prote
134	5	23.8	97	2	AH0965	probable secreted	207	5	23.8	157	2	A71137	hypothetical prote
135	5	23.8	98	2	T42307	hypothetical prote	208	5	23.8	157	2	S75191	hypothetical prote
136	5	23.8	99	2	AC0152	hypothetical prote	209	5	23.8	157	2	G97760	hypothetical prote
137	5	23.8	100	2	S14851	actin - fruit fly	210	5	23.8	159	2	E95328	hypothetical prote
138	5	23.8	100	2	A87463	conserved hypotet	211	5	23.8	159	2	G84125	hypothetical prote
139	5	23.8	101	2	I38611	calcium channel, v	212	5	23.8	161	2	S26039	actin - shore pine
140	5	23.8	101	2	H81352	small hydrophobic	213	5	23.8	161	2	B71687	hypothetical prote
141	5	23.8	102	2	S51974	probable membrane	214	5	23.8	165	2	PS0261	calcineurin regula
142	5	23.8	103	2	AB6584	Cr631 hypothetical	215	5	23.8	166	2	AE2392	hypothetical prote
143	5	23.8	104	2	S53942	probable membrane	216	5	23.8	166	2	AG3566	hypothetical prote
144	5	23.8	104	2	S51479	drought-induced pr	217	5	23.8	166	2	F97128	probable membrane
145	5	23.8	104	2	AG3518	hypothetical prote	218	5	23.8	167	2	B81676	probable signal pe
146	5	23.8	105	2	B81963	hypothetical IS110	219	5	23.8	167	2	B71553	hypothetical prote
147	5	23.8	106	2	H32537	T-cell receptor be	220	5	23.8	168	2	T25892	hypothetical prote
148	5	23.8	106	2	G97841	acetyl-CoA acetyl	221	5	23.8	170	1	A33391	calcineurin regula
149	5	23.8	107	2	D71114	hypothetical prote	222	5	23.8	170	1	S34127	calcineurin regula
150	5	23.8	107	2	JC1127	major allergen cha	223	5	23.8	170	1	JC1220	calcineurin regula
151	5	23.8	109	2	C56413	major allergen Fel	224	5	23.8	170	2	A35944	Ig gamma-2a chain
152	5	23.8	109	2	T35249	hypothetical prote	225	5	23.8	170	2	JC7242	calcineurin regula
153	5	23.8	113	2	JQ1128	gas-vesicle operon	226	5	23.8	171	2	T28000	hypothetical prote
154	5	23.8	113	2	T08234	gas-vesicle operon	227	5	23.8	173	2	C64161	hypothetical prote
155	5	23.8	113	2	F51025	transcription fact	228	5	23.8	173	2	H95399	probable transcrip
156	5	23.8	113	2	B81370	hypothetical prote	229	5	23.8	174	2	F82786	conserved hypotet
157	5	23.8	114	2	E71141	hypothetical prote	230	5	23.8	176	2	JQ1232	calcineurin regula
158	5	23.8	117	1	A59316	ghrelin precursor	231	5	23.8	177	2	B82221	conserved hypotet
159	5	23.8	117	2	S28124	gas-vesicle operon	232	5	23.8	177	2	AC2826	conserved hypotet
160	5	23.8	117	2	T49511	hypothetical prote	233	5	23.8	179	2	JC1221	calcineurin regula
161	5	23.8	118	2	T30509	hypothetical prote	234	5	23.8	180	2	C71869	hypothetical prote
162	5	23.8	119	2	S28133	gas vesicle protei	235	5	23.8	181	2	AB0127	Syd protein (impor
163	5	23.8	119	2	H70437	hypothetical prote	236	5	23.8	181	2	T05925	hypothetical prote
164	5	23.8	120	2	H84973	ribosome-binding f	237	5	23.8	182	2	B97604	hypothetical prote
165	5	23.8	120	2	F71104	hypothetical prote	238	5	23.8	183	2	G97436	lema protein (U661
166	5	23.8	120	2	B71020	hypothetical prote	239	5	23.8	183	2	AC2655	conserved hypotet
167	5	23.8	122	2	G81855	insertion element	240	5	23.8	183	2	AG0115	probable membrane
168	5	23.8	123	2	F83507	hypothetical prote	241	5	23.8	184	2	AB0207	probable exported
169	5	23.8	125	2	A26337	actin, skeletal mu	242	5	23.8	184	2	H87676	ATP synthase F1, d
170	5	23.8	125	2	C83138	hypothetical prote	243	5	23.8	184	2	D83234	hypothetical prote
171	5	23.8	128	2	AB2487	hypothetical prote	244	5	23.8	186	2	F96718	hypothetical prote
172	5	23.8	128	2	AE1701	hypothetical prote	245	5	23.8	187	2	AH1870	hypothetical prote
173	5	23.8	129	2	B90175	NADH dehydrogenase	246	5	23.8	187	2	G72748	hypothetical prote
174	5	23.8	130	2	PL0098	Ig heavy chain pre	247	5	23.8	188	2	A64332	probable phosphono
175	5	23.8	130	2	G71212	hypothetical prote	248	5	23.8	189	2	E57498	outer membrane ser

249	5	23.8	190	2	D75401	transcription repr	322	250	2	T30502	hypothetical prote
250	5	23.8	193	2	T60019	hypothetical prote	323	251	2	T45028	hypothetical prote
251	5	23.8	193	1	B69043	conserved hypotet	324	252	2	F83604	hypothetical prote
252	5	23.8	193	2	AF1981	hypothetical prote	325	252	2	T15300	hypothetical prote
253	5	23.8	193	2	AG3389	hypothetical prote	326	252	2	T31439	probable cohyr a
254	5	23.8	193	2	S39829	probable membrane	327	253	2	F87672	conserved hypotet
255	5	23.8	194	2	AI0241	hypothetical prote	328	254	2	B31790	ig heavy chain v r
256	5	23.8	195	2	S39777	actin beta - pig (329	256	2	E88469	protein C28H8.6 [1
257	5	23.8	195	2	T21997	hypothetical prote	330	257	2	G95252	pns system, IIC co
258	5	23.8	195	2	A99346	hypothetical prote	331	257	2	E98117	hypothetical prote
259	5	23.8	196	2	G65039	hypothetical prote	332	257	2	D70341	hypothetical prote
260	5	23.8	196	2	B71232	hypothetical prote	333	257	2	B96003	probable transcrip
261	5	23.8	197	2	D97061	uncharacterized lo	334	258	2	T32939	phosphatidylserine
262	5	23.8	200	2	B69168	hypothetical prote	335	259	2	G83813	hypothetical prote
263	5	23.8	200	2	D95950	probable esterase	336	262	2	A70428	hypothetical prote
264	5	23.8	202	2	G81997	probable phage ant	337	262	2	AI3570	nickel transport A
265	5	23.8	203	2	G81700	thymidylate kinase	338	265	2	E72592	hypothetical prote
266	5	23.8	204	2	S35256	actin (clone fac9)	339	265	2	S76508	hypothetical prote
267	5	23.8	206	2	A81231	50S ribosomal prot	340	265	2	D70778	hypothetical prote
268	5	23.8	207	2	D83187	hypothetical prote	341	265	2	B90242	conserved hypotet
269	5	23.8	207	2	AH1190	3-methyladenine DN	342	266	1	R5H07A	ribosomal protein
270	5	23.8	207	2	AH1548	3-methyladenine DN	343	266	1	R5RT7A	ribosomal protein
271	5	23.8	210	2	H82037	conserved hypotet	344	266	2	A69358	carboxylesterase (
272	5	23.8	210	2	B84121	stage II sporulati	345	266	2	A30241	ribosomal protein
273	5	23.8	211	2	T19859	hypothetical prote	346	266	2	T52089	ribosomal protein
274	5	23.8	212	2	B81669	conserved hypotet	347	266	2	I40626	probable sigma fac
275	5	23.8	213	2	S68213	conserved hypotet	348	266	2	A97109	sigma factor E pro
276	5	23.8	213	2	A61043	ig heavy chain (Ma	349	267	1	DEBTOP	crotonine-5'-phos
277	5	23.8	214	2	S17680	actin CA15 - sea s	350	267	2	AI3458	urease accessory p
278	5	23.8	214	2	PC4202	fibrinolytic prote	351	269	1	ICHU1B	interleukin-1 beta
279	5	23.8	214	2	C86967	monoclonal antibod	352	269	2	S77394	hypothetical prote
280	5	23.8	214	2	B96755	probable membrane	353	269	2	F97101	flagellin [impor
281	5	23.8	215	2	T01349	hypothetical prote	354	270	2	T48780	hypothetical prote
282	5	23.8	216	1	S42716	calcineurin regula	355	270	2	B81752	conserved hypotet
283	5	23.8	216	2	T30657	hypothetical prote	356	270	2	D83542	hypothetical prote
284	5	23.8	216	2	B82612	hypothetical prote	357	271	2	T19761	hypothetical prote
285	5	23.8	216	2	AE2348	hypothetical prote	358	271	2	F71490	hypothetical prote
286	5	23.8	217	1	H71208	hypothetical prote	359	271	2	AI3085	conserved hypotet
287	5	23.8	218	2	A70964	probable deda prot	360	271	2	G87440	hypothetical prote
288	5	23.8	218	2	T23318	hypothetical prote	361	272	2	G71296	probable spermidin
289	5	23.8	218	2	T23470	hypothetical prote	362	272	2	H71441	hypothetical prote
290	5	23.8	220	2	S68211	hypothetical prote	363	274	2	T31528	hypothetical prote
291	5	23.8	221	2	S49220	ig heavy chain (Ma	364	274	2	AE3552	high-affinity bran
292	5	23.8	222	2	E90400	ig gamma-1 chain -	365	275	2	S74417	hypothetical prote
293	5	23.8	224	2	H70749	hypothetical prote	366	276	2	S12962	hypothetical prote
294	5	23.8	226	2	S75665	ABC-type transport	367	277	2	AF2404	probable 2-oxoisov
295	5	23.8	226	2	T36096	probable secreted	368	278	2	C69506	hypothetical prote
296	5	23.8	228	2	G86847	hypothetical prote	369	280	2	T26582	conserved hypotet
297	5	23.8	230	2	T45555	hypothetical prote	370	280	2	H95901	conserved hypotet
298	5	23.8	231	2	PC4155	H+-transporting AT	371	281	2	D72351	hypothetical prote
299	5	23.8	231	2	AF0866	ig gamma-2b chain	372	281	2	AF2090	hypothetical prote
300	5	23.8	231	2	T50222	DNA mismatch repai	373	282	2	A48516	hypothetical prote
301	5	23.8	232	2	T46679	hypothetical lysin	374	283	2	AB0631	3,4-dihydroxypheny
302	5	23.8	232	2	T29440	alpha-smooth muscl	375	283	2	S61156	hypothetical prote
303	5	23.8	233	2	C71288	hypothetical prote	376	284	2	H64914	probable dimethyls
304	5	23.8	233	2	T22977	probable purine nu	377	284	2	H90915	probable DMSO redu
305	5	23.8	233	2	D97118	hypothetical prote	378	284	2	E85764	hypothetical prote
306	5	23.8	234	2	B84177	RNA polymerase sig	379	284	2	A99201	hypothetical prote
307	5	23.8	234	2	G75198	hypothetical prote	380	285	2	AB0215	3,4-dihydroxypheny
308	5	23.8	235	2	A97996	hypothetical prote	381	285	2	F70348	hypothetical prote
309	5	23.8	236	2	T59323	degenerate transpo	382	286	2	T22824	hypothetical prote
310	5	23.8	236	2	A77281	apolipoprotein B m	383	287	2	C84567	hypothetical prote
311	5	23.8	236	2	A75455	hypothetical prote	384	287	2	G70974	probable cmaal pro
312	5	23.8	239	1	BI4655	hypothetical prote	385	287	2	A70614	probable mma22 pro
313	5	23.8	240	2	D82591	virB1 protein prec	386	287	2	B70829	probable uma22 pro
314	5	23.8	241	2	D87494	protein phosphatas	387	287	2	AH3492	hydroxymethylgluta
315	5	23.8	242	2	D84422	conserved hypotet	388	288	2	A71148	hypothetical prote
316	5	23.8	243	1	S13368	60S ribosomal prot	389	290	2	G71304	probable translati
317	5	23.8	243	2	S39850	licheninase (EC 3	390	290	2	E97637	hypothetical prote
318	5	23.8	246	2	T00692	ig gamma chain - m	391	290	2	AF2860	xanthine dehydroge
319	5	23.8	247	1	F22MG	60S ribosomal prot	392	291	2	B69442	protein -coa lig
320	5	23.8	248	2	C89472	NADH dehydrogenase	393	291	2	I48602	insulin-like growt
321	5	23.8	248	2	D86676	protein ZC53.3 [im	394	291	2	T03900	hypothetical prote

395	5	23.8	292	2	T34997	lysR-type transcri	468	5	23.8	324	1	GIMS	Ig gamma-1 chain C
396	5	23.8	293	2	AF3647	phosphatidylglycer	469	5	23.8	324	2	T37228	probable sugar tra
397	5	23.8	294	2	T33825	hypothetical prote	470	5	23.8	324	2	T48236	hypothetical prote
398	5	23.8	295	2	AE1927	3-ketoacyl-acyl ca	471	5	23.8	325	2	S48698	3-dehydroquinat d
399	5	23.8	296	2	E75557	glutamine cyclotha	472	5	23.8	325	2	JC2008	actin homolog prot
400	5	23.8	297	2	A97337	conserved hypothet	473	5	23.8	325	2	T32940	hypothetical prote
401	5	23.8	298	2	A95069	leucine-rich prote	474	5	23.8	325	2	D87733	protein W03B8.9 [1
402	5	23.8	299	2	H81889	probable transposa	475	5	23.8	326	2	D71521	probable malate de
403	5	23.8	300	2	B83433	translocator prote	476	5	23.8	326	2	C81678	malate dehydrogena
404	5	23.8	301	2	H91117	probable transcrip	477	5	23.8	327	2	S11452	actin (clone 302)
405	5	23.8	302	2	H87216	Mycolic acid synth	478	5	23.8	327	2	AG3448	transcription regu
406	5	23.8	303	2	G85962	probable transcrip	479	5	23.8	328	2	J60375	carbonic anhydrase
407	5	23.8	304	2	G83216	probable transcrip	480	5	23.8	328	2	C64845	probable membrane
408	5	23.8	305	2	B83823	endonuclease IV BH	481	5	23.8	328	2	E85651	hypothetical prote
409	5	23.8	306	2	AC3185	transcription regu	482	5	23.8	328	2	B90791	hypothetical prote
410	5	23.8	307	2	B64050	glucose kinase hom	483	5	23.8	329	2	H75432	alcohol dehydrogen
411	5	23.8	308	2	B83554	probable transfe	484	5	23.8	329	2	AG1528	Ig gamma-2a chain
412	5	23.8	309	2	A37766	SEC14 protein - Ye	485	5	23.8	330	1	G2MSA	hypothetical prote
413	5	23.8	310	2	AF2223	heterodisulfide re	486	5	23.8	330	2	E2535	hypothetical prote
414	5	23.8	311	2	H70505	hypothetical prote	487	5	23.8	331	2	S24409	actin - brown alga
415	5	23.8	312	2	C64829	transcription regu	488	5	23.8	331	2	AE3372	toluene tolerance
416	5	23.8	313	2	B69903	aromatic metabolit	489	5	23.8	332	2	D84752	hypothetical prote
417	5	23.8	314	2	H37212	co/Zn/Cd efflux sy	490	5	23.8	332	2	T21279	hypothetical prote
418	5	23.8	315	2	C83967	aspartate carbamoy	491	5	23.8	333	2	S64337	transaldolase (EC
419	5	23.8	316	2	H95162	hypothetical prote	492	5	23.8	333	2	S75980	hypothetical prote
420	5	23.8	317	2	G98028	hypothetical prote	493	5	23.8	333	2	B72306	hypothetical prote
421	5	23.8	318	2	B64960	membrane protein Y	494	5	23.8	334	2	B95894	probable threonine
422	5	23.8	319	2	H90965	hypothetical prote	495	5	23.8	334	2	E81152	ADPglyceromanno-he
423	5	23.8	320	2	H85813	hypothetical prote	496	5	23.8	334	2	C81122	ADPglyceromanno-he
424	5	23.8	321	2	H81966	probable transcrip	497	5	23.8	334	2	S51650	probable iron/asco
425	5	23.8	322	2	H81022	transcription regu	498	5	23.8	334	2	AG2467	cyclin delta-1 - A
426	5	23.8	323	2	AF0315	probable transcrip	499	5	23.8	335	1	G2MSAB	hypothetical prote
427	5	23.8	324	2	F83254	conserved hypothet	500	5	23.8	335	2	D95119	Ig gamma-2a chain
428	5	23.8	325	2	S18533	eryG protein - Sac	501	5	23.8	335	2	G81877	iron-compound ABC
429	5	23.8	326	2	T19412	hypothetical prote	502	5	23.8	335	2	D81893	probable transposa
430	5	23.8	327	2	E82383	transcription regu	503	5	23.8	335	2	H97988	hypothetical prote
431	5	23.8	328	2	T27915	hypothetical prote	504	5	23.8	336	2	T04085	actin - maize (fra
432	5	23.8	329	2	E83129	homoprotocatechua	505	5	23.8	336	2	E69852	malate dehydrogena
433	5	23.8	330	2	D97605	proteinase chain h	506	5	23.8	337	2	S44207	hypothetical prote
434	5	23.8	331	2	AF2827	HFIC protein (limp	507	5	23.8	338	1	S44207	flagellar motor sw
435	5	23.8	332	2	A03000	actin 3 - fruit fl	508	5	23.8	338	2	G83508	hypothetical prote
436	5	23.8	333	2	AD0502	homoserine kinase	509	5	23.8	338	2	S45765	hypothetical prote
437	5	23.8	334	2	G75286	serine proteinase	510	5	23.8	339	2	C88035	protein M01D1.4 [1
438	5	23.8	335	2	AE3558	mRNA-binding prote	511	5	23.8	339	2	T19555	hypothetical prote
439	5	23.8	336	2	AG2409	modulation protein	512	5	23.8	339	2	T19600	probable transposi
440	5	23.8	337	2	A24400	ribosomal protein	513	5	23.8	339	2	T32676	hypothetical prote
441	5	23.8	338	2	S53836	hypothetical prote	514	5	23.8	339	2	T26766	hypothetical prote
442	5	23.8	339	2	T24625	two-component resp	515	5	23.8	339	2	T28781	hypothetical prote
443	5	23.8	340	2	C83989	transcription regu	516	5	23.8	339	2	T15113	hypothetical prote
444	5	23.8	341	2	OXBP2L	hypothetical prote	517	5	23.8	339	2	T24136	hypothetical prote
445	5	23.8	342	2	S11787	2-dehydro-3-deoxyg	518	5	23.8	339	2	T23362	hypothetical prote
446	5	23.8	343	2	AG0485	hypothetical prote	519	5	23.8	339	2	T20442	hypothetical prote
447	5	23.8	344	2	T29607	hypothetical prote	520	5	23.8	339	2	T15721	hypothetical prote
448	5	23.8	345	2	T18991	hypothetical prote	521	5	23.8	339	2	T18783	hypothetical prote
449	5	23.8	346	2	T29396	hypothetical prote	522	5	23.8	339	2	T26396	hypothetical prote
450	5	23.8	347	2	C95941	probable transcrip	523	5	23.8	339	2	T22780	hypothetical prote
451	5	23.8	348	2	C95941	ferrichrome ABC tr	524	5	23.8	339	2	T19261	hypothetical prote
452	5	23.8	349	2	B83778	probable transposa	525	5	23.8	339	2	T27715	hypothetical prote
453	5	23.8	350	2	H81916	conserved hypothet	526	5	23.8	339	2	T25335	hypothetical prote
454	5	23.8	351	2	H82907	hypothetical prote	527	5	23.8	339	2	T25365	hypothetical prote
455	5	23.8	352	2	T12516	capsular polysacch	528	5	23.8	340	2	A35009	thermopsin (EC 3.4
456	5	23.8	353	2	F70441	probable LysR-fam	529	5	23.8	340	2	C83124	probable transcrip
457	5	23.8	354	2	B95373	hypothetical prote	530	5	23.8	341	1	P3VVG	p3 protein - grape
458	5	23.8	355	2	E82520	hypothetical prote	531	5	23.8	341	2	G95888	hypothetical prote
459	5	23.8	356	2	F64384	hypothetical prote	532	5	23.8	341	2	AB1394	galactosyltransfer
460	5	23.8	357	2	AE3635	dihydrodipicolinat	533	5	23.8	341	2	AE1769	galactosyltransfer
461	5	23.8	358	2	E82420	transcription regu	534	5	23.8	342	2	C69644	ketol-acid reducto
462	5	23.8	359	2	T24948	hypothetical prote	535	5	23.8	342	2	T21098	protein kinase (EC
463	5	23.8	360	2	T27308	hypothetical prote	536	5	23.8	342	2	T51703	nicotinate-nucleot
464	5	23.8	361	2	D84567	hypothetical prote	537	5	23.8	343	2	T17417	probable transcrip
465	5	23.8	362	2	AE3527	transcription regu	538	5	23.8	343	2	T01432	secy protein homol
466	5	23.8	363	2	G83291	probable transcrip	539	5	23.8	344	1	KHPGD	cathepsin D (EC 3.
467	5	23.8	364	2			540	5	23.8	344	1		

541	5	23.8	344	2	T27410	hypothetical prote	614	370	2	A29664	actin - sea urchin
542	5	23.8	345	2	D83551	conserved hypothet	615	370	2	D70076	ethanolamine trans
543	5	23.8	347	2	S12955	calcium channel pr	616	370	2	F81133	IS1106 transposase
544	5	23.8	347	2	T47070	hypothetical prote	617	370	2	E82369	thiH protein VC006
545	5	23.8	347	2	H96504	hypothetical prote	618	371	2	E77345	hypothetical prote
546	5	23.8	347	2	T27315	hypothetical prote	619	371	2	F90486	dehydrogenase, pro
547	5	23.8	348	2	A57234	lin-44 protein pre	620	372	2	A75434	prephenate dehydro
548	5	23.8	348	2	T33544	hypothetical prote	621	372	2	B81952	probable type II r
549	5	23.8	348	2	T23430	hypothetical prote	622	373	2	S57262	actin - red alga (
550	5	23.8	349	2	B28819	actin, fetal skele	623	373	2	G81302	queuine tRNA-ribos
551	5	23.8	350	2	S06758	glycerol-3-phospha	624	373	2	S54545	hypothetical prote
552	5	23.8	350	2	A69834	conserved hypothet	625	374	1	ATBOB	actin beta - bovin
553	5	23.8	350	2	S34557	hypothetical prote	626	374	1	ATBOG	gamma-actin - huma
554	5	23.8	350	2	T50581	binding-protein-de	627	374	1	JCS818	gamma-actin - bovi
555	5	23.8	350	2	S61646	dolichyl-diphospho	628	374	1	A95960	probable cytochrom
556	5	23.8	350	2	T34002	hypothetical prote	629	375	1	ATBB	actin, skeletal mu
557	5	23.8	350	2	AE1845	hypothetical prote	630	375	1	ATBOSM	actin, aortic smoo
558	5	23.8	351	2	T40933	hypothetical prote	631	375	1	ATHUB	actin beta - mouse
559	5	23.8	351	2	B70811	hypothetical prote	632	375	1	ATMSB	actin beta - mouse
560	5	23.8	351	2	T51513	hypothetical prote	633	375	1	ATRTC	actin beta - rat
561	5	23.8	351	2	E97131	N-acetylmuramidase	634	375	1	ATRBB	actin beta - chick
562	5	23.8	352	2	JS0023	glycerol-3-phospha	635	375	1	ATCHB	actin beta, cytosk
563	5	23.8	352	2	A49210	heat shock protein	636	375	1	A48324	actin gamma 1 - hu
564	5	23.8	352	2	T2781	hypothetical prote	637	375	1	ATHUG	actin gamma - mous
565	5	23.8	353	1	K1B8V	thymidine kinase (638	375	1	ATHUG	actin gamma, cytos
566	5	23.8	353	2	S31790	glycerol-3-phospha	639	375	1	S11222	actin gamma, cytos
567	5	23.8	353	2	S06760	glycerol-3-phospha	640	375	1	ATDO	actin - slime mold
568	5	23.8	353	2	E84941	imidazoleglycerol-	641	375	1	ATAX	actin - maize
569	5	23.8	353	2	AF1606	aminopeptidase p h	642	375	1	ATZM1	actin - Acanthamo
570	5	23.8	353	2	B88485	sulfate ABC transp	643	375	2	S70377	actin - Phaffia rh
571	5	23.8	355	2	G84040	nitrogenase cofact	644	375	2	S71124	actin beta-1, cyto
572	5	23.8	355	2	S16047	hypothetical prote	645	375	2	S71125	actin beta-2, cyto
573	5	23.8	355	2	T20782	probable intercell	646	375	2	S71126	actin beta, cytosk
574	5	23.8	355	2	S77610	actin - Oxytricha	647	375	2	S42103	actin - Puccinia g
575	5	23.8	357	1	ATQO	hypothetical prote	648	375	2	A50001	actin beta - goose
576	5	23.8	357	2	C70805	hypothetical prote	649	375	2	S47897	actin 1 - Pneumocy
577	5	23.8	358	2	T37805	hypothetical prote	650	375	2	A26836	actin - fission ye
578	5	23.8	360	2	S06759	glycerol-3-phospha	651	375	2	JT0385	actin gamma - Emer
579	5	23.8	360	2	H82081	phospho-N-acetylmu	652	375	2	S03126	actin - imperfect
580	5	23.8	360	2	S65210	hypothetical prote	653	375	2	A31134	actin, macronuclea
581	5	23.8	361	1	D69005	hypothetical prote	654	375	2	A54728	actin alpha, cardi
582	5	23.8	361	2	S68089	actin 2 - Arabidop	655	375	2	S33866	actin, cytosolic (
583	5	23.8	362	2	S21963	glycerol-3-phospha	656	375	2	T25272	hypothetical prote
584	5	23.8	362	2	A26559	actin type 5, cyto	657	376	1	ATCHSM	actin gamma, smoot
585	5	23.8	362	2	S68090	actin 8 - Arabidop	658	376	1	ATURS	actin Cyl - sea ur
586	5	23.8	362	2	G72672	probable integrase	659	376	1	A43552	actin gamma, cytos
587	5	23.8	363	2	S23137	glycerol-3-phospha	660	376	1	ATFF7	actin 7 - fruit fl
588	5	23.8	363	2	T02522	origin recognition	661	376	1	ATFF8	actin 8 - fruit fl
589	5	23.8	364	2	AG2537	cysteine synthase	662	376	1	ATFY	actin - slime mold
590	5	23.8	364	2	D70164	heat shock protein	663	376	1	ATAXE	actin - Entamoeba
591	5	23.8	364	2	H72059	2-component sensor	664	376	1	ATR23	actin 3 - rice
592	5	23.8	364	2	E86563	2-component sensor	665	376	1	ATR27	actin 7 - rice
593	5	23.8	364	2	AC0589	UDPgalactopyranose	666	376	1	ATSK3	actin - soybean
594	5	23.8	365	1	FCCE82	translation releas	667	376	2	S55892	endo-1,4-beta-xyla
595	5	23.8	365	1	A36480	translation releas	668	376	2	S71123	actin alpha-anomal
596	5	23.8	365	1	A64190	translation releas	669	376	2	S12628	actin - malaria pa
597	5	23.8	365	2	AF1952	cysteine synthase	670	376	2	S07639	actin - yeast (Can
598	5	23.8	365	2	S49007	actin - Pythium ir	671	376	2	A54496	actin I - malaria
599	5	23.8	365	2	A37431	actin, type 1 - Em	672	376	2	S49480	actin 3 - Atlantic
600	5	23.8	365	2	G91099	peptide chain rele	673	376	2	JN0832	actin (clone gen3)
601	5	23.8	365	2	G85944	peptide chain rele	674	376	2	S49481	actin 5 - Atlantic
602	5	23.8	365	2	AC0871	peptide chain rele	675	376	2	JN0833	actin (clones Ia a
603	5	23.8	366	1	OORT52	bradykinin recepto	676	376	2	S43509	actin - California
604	5	23.8	366	2	G84849	probable actin (im	677	376	2	S49479	actin 11 - Atlanti
605	5	23.8	366	2	AB0109	peptide chain rele	678	376	2	A54509	actin 11 - malaria
606	5	23.8	366	2	T26624	hypothetical prote	679	376	2	B23412	actin 12 - slime m
607	5	23.8	367	2	JT0596	actin Ardd - slime	680	376	2	A25084	actin 15 - slime m
608	5	23.8	367	2	B88969	protein T27B7.1 (i	681	376	2	S11453	actin (clone 403)
609	5	23.8	367	2	T46021	hypothetical prote	682	376	2	JQ0154	actin - Hydra atte
610	5	23.8	368	2	S71150	heat shock protein	683	376	2	S07382	actin A2 - silkwor
611	5	23.8	369	2	T42295	hypothetical prote	684	376	2	S09059	actin A1 - silkwor
612	5	23.8	369	2	F83250	histidinol-phospha	685	376	2	S12730	actin - California
613	5	23.8	369	2	T22708	hypothetical prote	686	376	2	S04538	actin 87E - fruit

687	23.8	376	2	JS0190	actin, muscle - st	760	5	23.8	378	1	ATSY1	actin 1 - soybean
688	23.8	376	2	S16709	actin 2 - Caenorha	761	5	23.8	378	2	S21907	actin, muscle - se
689	23.8	376	2	S27135	actin 4 - Caenorha	762	5	23.8	378	2	A43911	actin 1-alpha, mus
690	23.8	376	2	C23412	actin 3-sub1 - sli	763	5	23.8	379	1	JC2142	alkaline proteinas
691	23.8	376	2	S11451	actin (clone 211)	764	5	23.8	379	1	ATR22	actin 2 - rice
692	23.8	376	2	S11450	actin (clone 205)	765	5	23.8	379	2	S33387	actin, muscle - se
693	23.8	376	2	S24408	actin - Achlya bis	766	5	23.8	379	2	T39400	actin, like protein
694	23.8	376	2	A45634	actin - Cryptospor	767	5	23.8	379	2	T26246	hypothetical prote
695	23.8	376	2	A29407	actin - Tetrahymen	768	5	23.8	379	2	H96548	unknown protein [i
696	23.8	376	2	S07284	actin - Tetrahymen	769	5	23.8	380	2	F81449	8-amino-7-oxonon
697	23.8	376	2	A48449	Actin-1A - nematod	770	5	23.8	380	2	B64533	cystathionine gamm
698	23.8	376	2	JC1246	actin - fruit fly	771	5	23.8	380	2	D71973	probable cystathio
699	23.8	376	2	JE0414	actin A - Phytopht	772	5	23.8	380	2	S07002	actin 1 - carrot
700	23.8	376	2	S16710	actin 1 and actin	773	5	23.8	381	2	S07003	actin 2 - carrot
701	23.8	376	2	A0189	actin, cytosolic -	774	5	23.8	381	2	G96804	hypothetical prote
702	23.8	376	2	A25135	actin A3, cytosoli	775	5	23.8	382	2	F96009	hypothetical membr
703	23.8	376	2	A32788	actin gamma, smoot	776	5	23.8	382	2	G83808	hypothetical prote
704	23.8	376	2	A31375	actin gamma, smoot	777	5	23.8	383	2	E96499	hypothetical prote
705	23.8	376	2	A0261	actin, smooth musc	778	5	23.8	383	2	S76812	hypothetical prote
706	23.8	376	2	S07288	actin gamma, enter	779	5	23.8	385	2	S66292	hypothetical prote
707	23.8	376	2	S09578	actin 15A - sea ur	780	5	23.8	385	2	H71960	hypothetical prote
708	23.8	376	2	S26435	actin - sea urchin	781	5	23.8	386	2	T06788	actin - garden pea
709	23.8	376	2	S25488	actin 1 - garden p	782	5	23.8	386	2	S76131	nitrogenase cofact
710	23.8	376	2	JC5227	actin 1 - earthwor	783	5	23.8	387	2	A45117	aspartic proteinas
711	23.8	376	2	JC5228	actin 2 - earthwor	784	5	23.8	387	2	A97709	hypothetical prote
712	23.8	376	2	JC5750	actin A4 - silkwor	785	5	23.8	388	2	C32905	desmoplakin, desmo
713	23.8	376	2	A44940	actin - pork tapew	786	5	23.8	389	2	S41748	heat shock protein
714	23.8	376	2	T24448	hypothetical prote	787	5	23.8	389	2	C82130	conserved hypothet
715	23.8	376	2	T24448	hypothetical prote	788	5	23.8	389	2	C90215	molybdopter in bios
716	23.8	376	2	T28935	actin, skeletal mu	789	5	23.8	389	2	H86656	hypothetical prote
717	23.8	377	1	ARHU	actin, skeletal mu	790	5	23.8	390	1	D64548	integral membrane
718	23.8	377	1	ARPT	actin, cardiac mus	791	5	23.8	390	1	RN2MS	polymerase-associa
719	23.8	377	1	A24904	actin 1 - rice	792	5	23.8	390	2	C95160	conserved hypothet
720	23.8	377	1	ATCH	actin 1 - sorghum	793	5	23.8	390	2	B84584	probable RING zinc
721	23.8	377	1	ARHUSM	actin - Cyanidiosc	794	5	23.8	390	2	G89877	conserved hypothet
722	23.8	377	1	A22224	actin alpha-1, ske	795	5	23.8	391	1	RN2YA	polymerase-associa
723	23.8	377	1	ATRBMS	actin alpha-2, aor	796	5	23.8	391	1	RN2ED	polymerase-associa
724	23.8	377	1	A25719	actin alpha, smoot	797	5	23.8	391	1	AB0443	hypothetical prote
725	23.8	377	1	ARHUC	actin alpha, aorti	798	5	23.8	392	2	E71633	acetyl-CoA acetyl
726	23.8	377	1	A3022	actin, cardiac mus	799	5	23.8	392	2	A84125	probable Grp-bind
727	23.8	377	1	AR621	actin, cardiac mus	800	5	23.8	392	2	T38450	probable Grp-bind
728	23.8	377	2	S60147	actin 1 - rice	801	5	23.8	393	1	GLMSM	lg gamma-1 chain C
729	23.8	377	2	S65079	actin - Cyanidiosc	802	5	23.8	393	2	T31121	hypothetical prote
730	23.8	377	2	S71118	actin alpha-1, ske	803	5	23.8	395	2	C82191	phospho-2-dehydro-
731	23.8	377	2	S71119	actin alpha-2, ske	804	5	23.8	395	2	T05680	hypothetical prote
732	23.8	377	2	S71120	actin alpha, cardi	805	5	23.8	395	2	S70912	related to ABC tra
733	23.8	377	2	S14120	actin - Volvox car	806	5	23.8	395	2	G96917	CMP-N-acetylneuram
734	23.8	377	2	S1933	actin - common tob	807	5	23.8	396	1	E64143	chloramphenicol re
735	23.8	377	2	B24848	actin alpha-3, ske	808	5	23.8	396	1	S39674	ywhd protein - Bac
736	23.8	377	2	A24848	actin alpha-1, car	809	5	23.8	396	2	T47207	aspartic proteinas
737	23.8	377	2	A29686	actin alpha-2, ske	810	5	23.8	396	2	S54641	probable membrane
738	23.8	377	2	S58316	actin - garden pea	811	5	23.8	396	2	T01201	hypothetical prote
739	23.8	377	2	JC4612	actin - Chlamydomo	812	5	23.8	396	2	AB2309	6-phosphofructo-2-
740	23.8	377	2	S20093	actin 101 - potato	813	5	23.8	397	2	S61066	hypothetical prote
741	23.8	377	2	S20098	actin 97 - potato	814	5	23.8	397	2	F90182	conserved hypothet
742	23.8	377	2	S20094	actin 58 - potato	815	5	23.8	397	2	B81223	hypothetical prote
743	23.8	377	2	S20096	actin 75 - potato	816	5	23.8	397	2	B81994	probable transmem
744	23.8	377	2	S68109	actin 11 - Arabido	817	5	23.8	397	2	B86295	hypothetical prote
745	23.8	377	2	S68112	actin 3 [imported]	818	5	23.8	398	2	T20901	hypothetical prote
746	23.8	377	2	S68110	actin 12 - Arabido	819	5	23.8	398	2	C91063	probable transpos
747	23.8	377	2	S68108	actin 4 - Arabidop	820	5	23.8	398	2	A57261	probable ISM5 tra
748	23.8	377	2	S68107	actin 7 - Arabidop	821	5	23.8	398	2	D95849	probable ISM5 tra
749	23.8	377	2	S20095	actin 71 - potato	822	5	23.8	399	1	G2MSAM	lg gamma-2a chain
750	23.8	377	2	JC5301	skeletal alpha-act	823	5	23.8	399	2	E86289	Ti6N11.7 protein -
751	23.8	377	2	T51180	actin [imported] -	824	5	23.8	400	2	JC4265	plasmalogen activa
752	23.8	377	2	T51181	actin [imported] -	825	5	23.8	400	2	T18060	hypothetical prote
753	23.8	377	2	T51182	actin [imported] -	826	5	23.8	401	2	T24381	hypothetical prote
754	23.8	377	2	T51183	actin isoform B [i	827	5	23.8	402	1	ITHUP1	plasmalogen activa
755	23.8	377	2	T51184	actin [imported] -	828	5	23.8	402	1	S06745	plasmalogen activa
756	23.8	377	2	T51185	actin [imported] -	829	5	23.8	402	2	D86575	phosphoglycerate k
757	23.8	377	2	T51176	actin [imported] -	830	5	23.8	402	2	C72049	phosphoglycerate k
758	23.8	377	2	T51177	actin [imported] -	831	5	23.8	402	2	G83289	conserved hypothet
759	23.8	377	2	T51178	actin ACT2 [import	832	5	23.8	404	2	B89819	pyrimidine nucleos

833	5	23.8	405	1	G2MSBM	Ig gamma-2b chain	906	5	23.8	442	2	D87351	conserved hypother
834	5	23.8	407	1	KHRVD	cathepsin D (EC 3.	907	5	23.8	443	1	S76611	hypothetical prote
835	5	23.8	408	1	VCWVSR	env polyprotein -	908	5	23.8	443	2	AE0826	probable cadaverin
836	5	23.8	409	2	B81417	hypothetical prote	909	5	23.8	443	2	H70430	K+ transport prote
837	5	23.8	410	2	B75010	hypothetical prote	910	5	23.8	443	2	S65963	flavastacin (EC 3.
838	5	23.8	410	1	KHMSD	cathepsin D (EC 3.	911	5	23.8	444	2	PC4436	monoclonal antibod
839	5	23.8	410	2	AD1080	arginine deiminase	912	5	23.8	444	2	A41842	lysine/cadaverine
840	5	23.8	410	2	T21960	hypothetical prote	913	5	23.8	444	2	B91268	transport of lysin
841	5	23.8	410	2	H95328	TamS transposase I	914	5	23.8	444	2	H86108	transport of lysin
842	5	23.8	411	1	I55504	platelet glycoprot	915	5	23.8	444	2	AC3162	conserved hypother
843	5	23.8	411	2	H64239	UV protection prot	916	5	23.8	445	2	G95204	conserved hypother
844	5	23.8	412	1	KHHUD	cathepsin D (EC 3.	917	5	23.8	446	2	S40295	Ig gamma-2a chain
845	5	23.8	412	2	T05285	farnesyl-diphospha	918	5	23.8	446	2	D89811	amidophosphoribos
846	5	23.8	412	2	T32577	hypothetical prote	919	5	23.8	447	2	B75212	hypothetical prote
847	5	23.8	413	2	T44924	farnesyl-diphospha	920	5	23.8	447	2	AE1026	hypothetical prote
848	5	23.8	413	2	E85408	probable squalene	921	5	23.8	447	2	H84797	hypothetical prote
849	5	23.8	413	2	AC0834	probable membrane	922	5	23.8	449	2	A71248	amidophosphoribos
850	5	23.8	413	2	AD0780	probable oxidoredu	923	5	23.8	449	2	F64536	3-deoxy-D-arabino-
851	5	23.8	415	2	B55563	AML2a protein - hu	924	5	23.8	449	2	D71971	phospho-2-dehydro-
852	5	23.8	415	2	S60078	Runt domain contai	925	5	23.8	449	2	D82215	phor histidine k
853	5	23.8	415	2	F96499	hypothetical prote	926	5	23.8	449	2	F97379	hypothetical prote
854	5	23.8	416	2	T32467	hypothetical prote	927	5	23.8	449	2	AE2597	conserved hypother
855	5	23.8	416	2	T44842	probable UDPglucos	928	5	23.8	450	2	B81730	sodium-dependent t
856	5	23.8	416	2	B81352	probable zinc prot	929	5	23.8	451	2	F69825	hypothetical prote
857	5	23.8	416	2	D75531	folyl-polyglutamat	930	5	23.8	451	2	T41911	hypothetical prote
858	5	23.8	416	2	B55649	TNFR-associated pr	931	5	23.8	451	2	T49095	beta-ketoacyl-CoA
859	5	23.8	416	2	S75793	poly(A) polymerase	932	5	23.8	451	2	T31790	hypothetical prote
860	5	23.8	417	2	F86731	exodeoxyribonuclea	933	5	23.8	452	2	F91292	hypothetical prote
861	5	23.8	417	2	T33207	hypothetical prote	934	5	23.8	452	2	A86134	hypothetical prote
862	5	23.8	417	2	AD0907	cytosine permease	935	5	23.8	453	2	F90327	cytosine permease
863	5	23.8	418	2	JC4069	ribX protein - Shi	936	5	23.8	453	2	B81409	probable transmemb
864	5	23.8	420	2	F84142	glutamate dehydrog	937	5	23.8	459	2	T45576	anthranilate N-hyd
865	5	23.8	421	2	E72573	probable 3-hydroxy	938	5	23.8	461	2	G83602	hypothetical prote
866	5	23.8	421	2	D95975	hypothetical outer	939	5	23.8	461	2	A46394	suppressor protein
867	5	23.8	422	2	T09742	drought-induced pr	940	5	23.8	463	1	P2WL13	L2 protein - humau
868	5	23.8	423	2	F75635	hypothetical prote	941	5	23.8	463	2	B83156	conserved hypother
869	5	23.8	424	1	SYECYT	tyrosine--trna lig	942	5	23.8	463	2	T36810	probable integral
870	5	23.8	424	2	B90922	tyrosine--trna synth	943	5	23.8	464	1	MNVUWC	nonstructural prot
871	5	23.8	424	2	AE0893	tyrosyl-trna synth	944	5	23.8	464	2	AG1146	beta-glucosidase h
872	5	23.8	426	2	T45767	hypothetical prote	945	5	23.8	464	2	AG1505	beta-glucosidase h
873	5	23.8	427	2	T45915	actin (ACT3) - Ara	946	5	23.8	464	2	G65071	hypothetical prote
874	5	23.8	427	2	B86408	probable dnaJ prot	947	5	23.8	464	2	D85943	probable proteogly
875	5	23.8	427	2	T39430	mitochondrial impo	948	5	23.8	464	2	T21505	hypothetical prote
876	5	23.8	428	2	AC3593	dihydroorotase (EC	949	5	23.8	465	2	D69785	beta-glucosidase h
877	5	23.8	428	2	F85770	tyrosine tRNA synt	950	5	23.8	466	2	F84139	beta-glucosidase h
878	5	23.8	430	1	AD0085	probable type II s	951	5	23.8	466	2	T44650	capsular polysacch
879	5	23.8	431	2	S32578	malC protein - Str	952	5	23.8	466	2	T08914	hypothetical prote
880	5	23.8	431	2	B84452	ctxX protein - Brw	953	5	23.8	466	2	F95307	conserved hypother
881	5	23.8	431	2	JN0317	hypothetical prote	954	5	23.8	467	1	FWBSBF	H+-transporting tw
882	5	23.8	432	1	AE0893	protein-tyrosine-p	955	5	23.8	467	1	MNVUW1	nonstructural prot
883	5	23.8	432	1	A34845	protein-tyrosine-p	956	5	23.8	468	2	T03164	probable major env
884	5	23.8	432	2	A25483	env polyprotein, r	957	5	23.8	468	2	S61964	probable membrane
885	5	23.8	432	2	H64152	hypothetical prote	958	5	23.8	469	2	AC3650	6-phosphogluconate
886	5	23.8	433	2	C64605	UDP-N-acetylglucos	959	5	23.8	469	2	D86144	protein probable U
887	5	23.8	433	2	A71909	udp-n-acetylglucos	960	5	23.8	469	2	S37483	Ig gamma-2a chain
888	5	23.8	435	1	TPHUN1	protein-tyrosine-p	961	5	23.8	472	1	A53048	translation initia
889	5	23.8	435	2	G95246	maltodextrin ABC t	962	5	23.8	474	1	G2MS11	Ig gamma-2b chain
890	5	23.8	435	2	D98111	hypothetical prote	963	5	23.8	474	2	T37967	4-aminobutyrate tr
891	5	23.8	435	2	B82848	alpha-ketoglutarat	964	5	23.8	474	2	A35732	protective protein
892	5	23.8	436	2	B84329	glutamyl-trna redu	965	5	23.8	474	2	T20108	hypothetical prote
893	5	23.8	436	2	B96828	probable aspartyl	966	5	23.8	475	1	S46941	translation initia
894	5	23.8	436	2	S77639	exopolysaccharide	967	5	23.8	475	2	S01321	Ig gamma-2b chain
895	5	23.8	438	2	T03513	probable secretion	968	5	23.8	478	2	H71224	hypothetical prote
896	5	23.8	439	2	E75118	probable threonine	969	5	23.8	478	2	T31955	hypothetical prote
897	5	23.8	440	2	T31458	protechlorophyllid	970	5	23.8	479	2	S61768	origin recognition
898	5	23.8	440	2	H71013	probable threonine	971	5	23.8	480	2	A31589	carboxypeptidase C
899	5	23.8	440	2	D95029	PIS system, IIC co	972	5	23.8	480	2	H82506	hypothetical prote
900	5	23.8	441	2	G97900	hypothetical prote	973	5	23.8	480	2	F82796	hypothetical prote
901	5	23.8	441	2	T01613	hypothetical prote	974	5	23.8	481	2	B84700	probable flavonol
902	5	23.8	441	2	F86185	hypothetical prote	975	5	23.8	481	2	E83062	deoxyribodipyrimid
903	5	23.8	442	2	H91097	hypothetical prote	976	5	23.8	481	2	G75192	probable ss-DNA-sp
904	5	23.8	442	2	H22209	hypothetical prote	977	5	23.8	481	2	T27435	hypothetical prote
905	5	23.8	442	2	S72568	hypothetical prote	978	5	23.8	482	1	A26934	nitrogen assimilat

979 5 23.8 482 2 F86902 dextranucrase (PC
 980 5 23.8 483 2 G86902 dextranucrase (PC
 981 5 23.8 484 2 H59298 hypothetical prote
 982 5 23.8 485 2 S4708 hypothetical prote
 983 5 23.8 486 2 S34800 sulfate transport
 984 5 23.8 487 2 A50310 NADH dehydrogenase
 985 5 23.8 488 2 T31296 butyrophilin - mou
 986 5 23.8 489 2 T31296 glycolate oxidase
 987 5 23.8 490 2 C71541 probable na-depend
 988 5 23.8 491 2 E59066 conserved hypothet
 989 5 23.8 492 2 H83286 probable chemotaxi
 990 5 23.8 493 2 B97221 L-fucose isomerase
 991 5 23.8 494 2 I40455 penicillin binding
 992 5 23.8 495 2 S46298 catalase (PC 1.11.
 993 5 23.8 496 2 S49147 BRP-2 protein - hu
 994 5 23.8 497 2 T25357 hypothetical prote
 995 5 23.8 498 2 D64578 conserved hypothet
 996 5 23.8 499 1 J00197 4-aminobutyrate tr
 997 5 23.8 498 2 T39077 hypothetical prote
 998 5 23.8 499 2 F83329 probable outer mem
 999 5 23.8 498 2 S28059 serum response fac
 1000 5 23.8 499 2 S70113 hypothetical prote

ALIGNMENTS

RESULT 1

S46737 hypothetical protein YHR038w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein H8179.10
 C:Species: Saccharomyces cerevisiae
 C>Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 23-Mar-2001
 C:Accession: S46737
 R:Du, Z.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of S. cerevisiae cosmid 8179.

A:Reference number: S46732

A:Accession: S46737

A:Molecule type: DNA

A:Residues: 1-230 <DUZ>

A:Cross-references: EMBL:U00062; NID:9488162; PID:9488171; GSPDB:GN00008; MIPS:YHR038w

C:Genetics:

A:Gene: MIPS:YHR038w

A:Map position: 8R

Query Match 33.3%; Score 7; DB 2; Length 230;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 RLFNRSF 15

Db 17 RLFNRSF 23

RESULT 2

S70367 stem cell factor short form precursor - quail
 C:Species: Coturnix coturnix (quail)
 C>Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000
 C:Accession: S70367
 R:Petitje, J.N.; Kulik, M.J.
 Biochim. Biophys. Acta 1307, 149-151, 1996

A:Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell fac

A:Reference number: S70366; MUID:96283808

A:Accession: S70367

A:Molecule type: mRNA

A:Residues: 1-253 <PET>

A:Cross-references: EMBL:U43079; NID:g1150877; PID:AAC59934.1; PID:g1150878

C:Superfamily: mouse mast cell growth factor

C:Keywords: growth factor; transmembrane protein

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-253/Product: stem cell factor short form #status predicted <MAT>

F:192-216/Domain: transmembrane #status predicted <TM>

Query Match 33.3%; Score 7; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. No. 4.1;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 13

Db 144 FFRLENR 150

RESULT 3

S70366 stem cell factor long form precursor - quail

C:Species: Coturnix coturnix (quail)

C>Date: 06-Dec-1996 #sequence_revision 25-Apr-1997 #text_change 21-Jul-2000

C:Accession: S70366

R:Petitje, J.N.; Kulik, M.J.

Biochim. Biophys. Acta 1307, 149-151, 1996

A:Title: Cloning and characterization of cDNAs encoding two forms of avian stem cell

A:Reference number: S70366; MUID:96283808

A:Accession: S70366

A:Molecule type: mRNA

A:Residues: 1-287 <PET>

A:Cross-references: EMBL:U43078; NID:g1150875; PID:AAC59933.1; PID:g1150876

C:Superfamily: mouse mast cell growth factor

C:Keywords: growth factor; transmembrane protein

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-287/Product: stem cell factor long form #status predicted <MAT>

F:226-250/Domain: transmembrane #status predicted <TM>

Query Match 33.3%; Score 7; DB 2; Length 287;

Best Local Similarity 100.0%; Pred. No. 4.6;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 13

Db 144 FFRLENR 150

RESULT 4

AD2284 hypothetical protein air3827 [imported] - Anabaena sp. (strain PCC 7120)

C:Species: Anabaena sp.

A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002

C:Accession: AD2284

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tasuda, M.; Tabata

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AD2284

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-472 <KUR>

A:Cross-references: GB:BA000019; PID:BA075526.1; PID:gl7132961; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: air3827

C:Superfamily: Synecocystis hypothetical protein sil0827

Query Match 33.3%; Score 7; DB 2; Length 472;

Best Local Similarity 100.0%; Pred. No. 7.2;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SFTQALG 20

Db 408 SFTQALG 414

RESULT 5

AI0479

probable membrane protein YPO3943 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C:Accession: A10479
 R:Parkhill, J.; Wren, B.W.; Thomson, M.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: A10479
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-565 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC93405.1; PID:g15981851; GSPDB:GN00175
 C:Genetics:
 A:Gene: YPO3943
 C:Superfamily: hypothetical protein b2380

Query Match 33.3%; Score 7; DB 2; Length 565;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLSTFF 9
 |||||
 Db 410 SLSTFF 416

RESULT 6
 A97416
 hypothetical protein AGR_C_830 [imported] - Agrobacterium tumefaciens (strain C58, Cere
 C:Species: Agrobacterium tumefaciens
 C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
 C:Accession: A97416
 R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, Science 294, 2323-2328, 2001
 A:Title: Genome sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A:Reference number: A97359; PMID:11743194
 A:Accession: A97416
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-43 <KUR>
 A:Cross-references: GB:AE007869; PIDN:AAK86282.1; PID:g15155394; GSPDB:GN00169
 C:Genetics:
 A:Gene: AGR_C_830
 A:Map position: circular chromosome

Query Match 28.6%; Score 6; DB 2; Length 43;
 Best Local Similarity 100.0%; Pred. No. 9.6;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSLSSTF 7
 |||||
 Db 31 GSLSSTF 36

RESULT 7
 T18094
 hypothetical protein A592R - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T18094
 R:Graves, M.V.; Van Etten, J.L.
 A:Reference number: Z18806
 A:Accession: T18094
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-69 <GRA>
 A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96934.1
 A:Experimental source: specific host Chlorella strain NC64A
 C:Genetics:

A:Note: A592R

Query Match 28.6%; Score 6; DB 2; Length 69;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLSTFF 8
 |||||
 Db 27 SLSTFF 32

RESULT 8
 B75387
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: B75387
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896
 A:Accession: B75387
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-122 <WHI>
 A:Cross-references: GB:AE001995; GB:AE000513; NID:g6459271; PIDN:AAF11085.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1518
 A:Map position: 1
 C:Superfamily: Deinococcus radiodurans hypothetical protein DR1518

Query Match 28.6%; Score 6; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLST 6
 |||||
 Db 66 SGSLST 71

RESULT 9
 AE1330
 hypothetical protein lmo2045 [imported] - Listeria monocytogenes (strain EGD-e)
 C:Species: Listeria monocytogenes
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AE1330
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
 ; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi,
 D.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.;
 ok, C.; Schluter, T.; Simoes, N.; Fierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla
 A:Title: Comparative genomics of Listeria species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AE1330
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-128 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00123.1; PID:g16411515; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo2045

Query Match 28.6%; Score 6; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
 |||||
 Db 60 LFNRSF 65

```

RESULT 10
TI7655
hypothetical protein al64L - Chlorella virus PBCV-1
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: TI7655
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: TI7655
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-134 <GRA>
A:Cross-references: EMBL:U42580; NID:q4028896; PIDN:AAC96532.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: al64L

Query Match 28.6%; Score 6; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 FNRSTF 16
|||||
Db 111 FNRSTF 116

RESULT 11
B70370
hypothetical protein aq_805 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: B70370
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Oviatt,
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666
A:Accession: B70370
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-155 <NOF>
A:Cross-references: GB:AE000708; NID:g2983356; PIDN:AA06954.1; PID:g2983367; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_805
C:Superfamily: Aquifex aeolicus hypothetical protein aq_805

Query Match 28.6%; Score 6; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
|||||
Db 54 LFNRSF 59

RESULT 12
S78572
ESR3 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YIL009c-a
C:Species: Saccharomyces cerevisiae
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 04-Mar-2000
C:Accession: S78572
R:Rowley, N.
submitted to the EMBL Data Library, August 1994
A:Reference number: S48442
A:Accession: S78572
A:Molecule type: DNA
A:Residues: 1-181 <ROW>
A:Cross-references: EMBL:Z38113; GSPDB:GN000009; MIPS:YIL009c-a

C:Genetics:
A:Gene: EST3; MIPS:YIL009c-a
A:Cross-references: MIPS:YIL009c-a
A:Map position: 9L
A:Introns: 92/3
C:Function:
A:Description: required for telomerase function
C:Superfamily: Saccharomyces cerevisiae EST3 protein

Query Match 28.6%; Score 6; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GSLSTF 7
|||||
Db 169 GSLSTF 174

RESULT 13
A56956
GTP-binding protein-associated protein (clone B) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: A56956
R:Janoueix-Lerosey, I.; Jollivet, F.; Camonis, J.; Marche, P.N.; Goud, B.
J. Biol. Chem. 270, 14801-14808, 1995
A:Title: Two-hybrid system screen with the small GTP-binding protein Rab6. Identifica
A:Reference number: A56956; MUID:95301579
A:Accession: A56956
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-186 <JAN>
A:Cross-references: GB:L40894; NID:g725273; PIDN:AAA78787.1; PID:g725274

Query Match 28.6%; Score 6; DB 2; Length 186;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLST 6
|||||
Db 130 SGSLST 135

RESULT 14
A44363
voltage-gated dihydropyridine-sensitive calcium channel subtype homolog (internal rep
C:Species: Homo sapiens (man)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 11-Jan-2000
C:Accession: A44363
R:Sun, W.; McPherson, J.D.; Hoang, D.O.; Wasmuth, J.J.; Evans, G.A.; Montal, M.
Genomics 14, 1092-1094, 1992
A:Title: Mapping of a human brain voltage-gated calcium channel to human chromosome 1
A:Reference number: A44363; MUID:93122776
A:Accession: A44363
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-197 <SUN>
A:Experimental source: brain
A:Note: sequence extracted from NCBI backbone (NCBI:122411, NCBIP:122412)
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 28.6%; Score 6; DB 2; Length 197;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
|||||
Db 150 TFFRLF 155

RESULT 15
H81731
CDP-diacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase TC0180 [imported]

```

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Jun-2000
 C:Accession: H81731
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
 C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
 A:Reference number: A81500; MUID:20150255
 A:Accession: H81731
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-202 <TET>
 A:Cross-references: GB:AE002285; GB:AE002160; MID:g7190212; PIDN:AAF39054.1; PID:g719021
 A:Experimental source: strain N199 (MoPn)
 C:Genetics:
 A:Gene: TC0180
 C:Superfamily: CDPdiacylglycerol--glycerol-3-phosphate 3-phosphatidyltransferase

Query Match 28.6%; Score 6; DB 2; Length 202;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TFFRLF 11
 DB 10 TFFRLF 15

Search completed: November 5, 2002, 11:07:10
 Job time : 21.1493 secs

1000

1000

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:58:02 ; Search time 5.64179 Seconds
(without alignments)
144.123 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 21

Sequence: 1 SGLSIFRFLNRSFTQALGK 21

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	33.3	230	1	FIL1_YEAST
2	7	33.3	500	1	ABF1_TRIKE
3	7	33.3	500	1	XYL1_TRIKO
4	6	28.6	155	1	Y805_AQUAE
5	6	28.6	169	1	CCAC_CAVPO
6	6	28.6	177	1	CCAC_CHICK
7	6	28.6	214	1	Y540_AQUAE
8	6	28.6	281	1	CCAD_MOUSE
9	6	28.6	281	1	CCAC_CHICK
10	6	28.6	287	1	SCF_CHICK
11	6	28.6	296	1	CEBB_MOUSE
12	6	28.6	297	1	CEBB_RAT
13	6	28.6	329	1	ISPB_HAEIN
14	6	28.6	345	1	CEBB_HUMAN
15	6	28.6	348	1	CEBB_BOVIN
16	6	28.6	354	1	FXFL_HUMAN
17	6	28.6	373	1	ACT6_DIPDE
18	6	28.6	405	1	WCAC_ECOLI
19	6	28.6	499	1	ABFB_ASPNG
20	6	28.6	561	1	YEHU_ECOLI
21	6	28.6	634	1	SL56_RAT
22	6	28.6	635	1	SL56_HUMAN
23	6	28.6	636	1	SL56_RABIT
24	6	28.6	637	1	MUTA_PROPR
25	6	28.6	699	1	YQOA_CAEEL
26	6	28.6	759	1	PMT6_YEAST
27	6	28.6	841	1	YAGX_ECOLI
28	6	28.6	970	1	Y087_BUCAL
29	6	28.6	1075	1	PST2_SCHPO
30	6	28.6	1126	1	RCOB_SULAC
31	6	28.6	1610	1	CCAD_MESAU
32	6	28.6	1787	1	UVRA_CHLMU
33	6	28.6	1807	1	ITB4_RAT

34	6	28.6	1852	1	CCAS_CYPCA
35	6	28.6	1966	1	CCAF_HUMAN
36	6	28.6	2139	1	CCAC_MOUSE
37	6	28.6	2161	1	CCAD_HUMAN
38	6	28.6	2169	1	CCAC_RAT
39	6	28.6	2171	1	CCAC_RABIT
40	6	28.6	2190	1	CCAD_CHICK
41	6	28.6	2203	1	CCAD_RAT
42	6	28.6	2221	1	CCAC_HUMAN
43	5	23.8	33	1	ACT_DICVI
44	5	23.8	67	1	YF82_ARCFU
45	5	23.8	87	1	CYSD_CHLFT
46	5	23.8	88	1	GP42_BSPPL
47	5	23.8	94	1	VAL5_VACCC
48	5	23.8	94	1	VAL5_VARV
49	5	23.8	97	1	YCA0_METJA
50	5	23.8	102	1	YAE5_YEAST
51	5	23.8	104	1	YGY9_YEAST
52	5	23.8	109	1	FEL2_FELCA
53	5	23.8	110	1	VLI_FPVVL
54	5	23.8	113	1	GVK1_HALNI
55	5	23.8	113	1	Y959_CAMJE
56	5	23.8	117	1	GHRL_CANFA
57	5	23.8	117	1	GHRL_HUMAN
58	5	23.8	117	1	GVPK_HALME
59	5	23.8	119	1	GVK2_HALNI
60	5	23.8	120	1	RBEA_BUCAL
61	5	23.8	120	1	YEF5_PYRHO
62	5	23.8	125	1	ACTS_PLEWA
63	5	23.8	128	1	IF2G_RABIT
64	5	23.8	130	1	FABP_BIOTA
65	5	23.8	130	1	GLOW_HYACE
66	5	23.8	131	1	FABP_LEPDS
67	5	23.8	139	1	A85A_MYCMR
68	5	23.8	140	1	Y337_MYCPN
69	5	23.8	140	1	YC49_ARCFU
70	5	23.8	141	1	YEF5_YEAST
71	5	23.8	148	1	YEAL_ECOLI
72	5	23.8	151	1	SP2C_STRPU
73	5	23.8	153	1	ACT4_LYPTI
74	5	23.8	153	1	Y359_RICPR
75	5	23.8	161	1	ACT_PINCO
76	5	23.8	161	1	Y311_RICPR
77	5	23.8	167	1	LSPA_CHLMU
78	5	23.8	169	1	CALB_HUMAN
79	5	23.8	169	1	CALB_MOUSE
80	5	23.8	172	1	ACT3_LYPTI
81	5	23.8	172	1	ACTM_LYPTI
82	5	23.8	173	1	YFHC_HAEIN
83	5	23.8	175	1	CALC_RAT
84	5	23.8	177	1	ATPD_BUCAP
85	5	23.8	178	1	CALC_MOUSE
86	5	23.8	180	1	ARF1_PLAFO
87	5	23.8	180	1	ARF1_PLAFO
88	5	23.8	181	1	Y65L_HORVU
89	5	23.8	188	1	YBF9_YEAST
90	5	23.8	193	1	KTHV_CHLMU
91	5	23.8	203	1	CADH_MOUSE
92	5	23.8	204	1	FADD_MOUSE
93	5	23.8	205	1	CADH_MOUSE
94	5	23.8	207	1	3MGH_LISMO
95	5	23.8	207	1	ACH2_LONAC
96	5	23.8	214	1	ACH2_LONAC
97	5	23.8	214	1	YQ37_MYCLE
98	5	23.8	218	1	YQ37_MYCTU
99	5	23.8	222	1	RISB_SPIOL
100	5	23.8	224	1	Y088_MYCTU
101	5	23.8	233	1	DEOD_TREPA
102	5	23.8	233	1	YSR3_CAEEL
103	5	23.8	236	1	ABME_HUMAN
104	5	23.8	239	1	VIB1_AGRTN9
105	5	23.8	243	1	GUB_BACLI
106	5	23.8	247	1	RL7_ARATH

P22316	cyprinus ca
O60840	homo sapien
O01815	mus musculus
O01668	homo sapien
P22002	rattus norv
P15381	oryctolagus
O73700	gallus gall
P27732	rattus norv
Q13936	homo sapien
Q24733	dictyocaulu
O28690	archaeoglob
P20958	chlorobium
O48398	bacterioph
P20992	vaccinia vi
P33840	variola vir
O58637	methanococc
P39725	saccharomyc
P33069	saccharomyc
P30440	felis silve
P06456	avian papil
P24375	halobacteri
Q9pux6	campylobact
Q9bef8	canis famil
Q9ubn3	homo sapien
O02236	halobacteri
P33963	halobacteri
P57457	buchnera ap
O59124	pyrococcus
P10994	pleurodeles
P33887	oryctolagus
Q17284	blomia trop
P81048	hyalophora
Q9u5p1	lepidoglyph
Q9kh57	mycobacteri
P75297	mycoplasma
O29019	archaeoglob
P32616	saccharomyc
P76240	escherichia
P04111	strongyloce
Q25380	lytechinus
Q9rdg9	rickettsia
P24902	pinus conto
Q9zd16	rickettsia
Q9PJY8	chlamydia m
P06705	homo sapien
Q63810	mus musculus
Q25379	lytechinus
Q25381	lytechinus
P44931	haemophilus
P28470	rattus norv
O51875	buchnera ap
Q63811	mus musculus
Q25761	plasmodium
Q94650	plasmodium
O48609	hordium vul
P58416	methanococc
P34224	saccharomyc
Q9pkx5	chlamydia m
Q70354	mus musculus
Q61160	mus musculus
Q92d89	listeria in
P58621	listeria mo
P23605	lonomia ach
Q49642	mycobacteri
P71936	mycobacteri
Q9XN32	spinacia oi
O10885	mycobacteri
O83716	treponema p
Q09951	caenorhabdi
P41238	homo sapien
P05350	agrobacteri
P27051	bacillus li
Q42208	arabidopsis

107	5	23.8	248	1	NUKC_MAIZE	P06670	zea mays (m	180	5	23.8	345	1	CATD_PTG	P00795	sus scrofa
108	5	23.8	252	1	YK6_CAREL	P41997	caenorhabdi	181	5	23.8	347	1	Y4RE_RHIN	P55525	rhizobium s
109	5	23.8	256	1	Y96_CAREL	O09476	caenorhabdi	182	5	23.8	349	1	GPDA_DROE	O27556	drosophila
110	5	23.8	257	1	Y453_AQUAE	O06760	aquifex aeo	183	5	23.8	349	1	GPDA_DROE	O27567	drosophila
111	5	23.8	262	1	ASH2_MOUSE	O35885	mus musculus	184	5	23.8	349	1	ISPG_CLOPE	P58667	clostridium
112	5	23.8	265	1	RL7A_FUGRU	O57592	fugu rubrip	185	5	23.8	350	1	OSTG_YEAST	P48439	saccharomyc
113	5	23.8	265	1	RL7A_HUMAN	P11518	homo sapien	186	5	23.8	350	1	YXGA_EUGGR	P15611	euglena gra
114	5	23.8	265	1	RL7A_MOUSE	P12970	mus musculus	187	5	23.8	352	1	GPDA_DROVI	O27735	drosophila
115	5	23.8	265	1	YK40_MYCTU	Q10522	mycobacteri	188	5	23.8	352	1	KITH_HSVMD	P17653	marek's dis
116	5	23.8	266	1	SP2G_CLOAB	Q45832	clostridium	189	5	23.8	353	1	ACT_ACECL	P53491	acetabulari
117	5	23.8	269	1	AROE_NEICI	P95337	neisseria c	190	5	23.8	353	1	GPDA_DROPS	Q27928	drosophila
118	5	23.8	269	1	IL1B_HUMAN	P01584	homo sapien	191	5	23.8	353	1	HIS7_BUCAL	P57203	buchnera ap
119	5	23.8	275	1	HPCB_ECOLI	Q05353	escherichia	192	5	23.8	353	1	SB11_CAEEL	P46506	caenorhabdi
120	5	23.8	278	1	ACT_CALFI	Q92192	calanus fin	193	5	23.8	355	1	NIFS_LACED	P31672	lactobacill
121	5	23.8	281	1	NAEL_THEMA	Q9wz53	thermologa	194	5	23.8	357	1	ACT1_OXYFA	P03503	oxytricha f
122	5	23.8	284	1	YNFH_ECGLI	P76173	escherichia	195	5	23.8	357	1	ACT2_OXYFA	P02583	oxytricha f
123	5	23.8	285	1	Y539_AQUAE	O66818	aquifex aeo	196	5	23.8	357	1	ACT2_OXYFA	P30172	solanum tub
124	5	23.8	287	1	CFAL_MYCTU	Q11195	mycobacteri	197	5	23.8	359	1	GPDA_DROKA	O97463	drosophila
125	5	23.8	290	1	EFTS_TREPA	O83614	treponema p	198	5	23.8	363	1	GPDA_DROME	P13706	drosophila
126	5	23.8	291	1	IBP3_MOUSE	P47878	mus musculus	199	5	23.8	363	1	ORC2_ARATH	Q38899	arabidopsis
127	5	23.8	291	1	SUD1_ARCFU	O28733	archaeoglob	200	5	23.8	364	1	DNAJ_BORBU	P28616	borrelia bu
128	5	23.8	298	1	END4_BACHD	Q9k433	bacillus ha	201	5	23.8	365	1	CATD_SHEEP	Q9mz58	ovis aries
129	5	23.8	300	1	PHAG_PSEAE	O51553	pseudomonas	202	5	23.8	365	1	RF2_ECGLI	P07012	escherichia
130	5	23.8	301	1	HGCI_HAETN	P44541	haemophilus	203	5	23.8	365	1	RF2_HAETN	P43918	haemophilus
131	5	23.8	302	1	YCAN_ECGLI	P24859	kluyveromyc	204	5	23.8	365	1	RF2_SALTY	P28353	saemonella
132	5	23.8	302	1	YCAN_ECGLI	P75836	escherichia	205	5	23.8	366	1	ACTD_PHYPO	P24263	physarum po
133	5	23.8	305	1	YFIB_BACHD	Q9k9v6	bacillus ha	206	5	23.8	370	1	YXER_BACSU	P24957	bacillus su
134	5	23.8	305	1	YEDI_ECGLI	P46125	escherichia	207	5	23.8	371	1	ACT2_NAEFO	P27132	naegleria f
135	5	23.8	307	1	CIW8_MOUSE	Q92211	mus musculus	208	5	23.8	371	1	ACT5_DIPDE	P53458	diphylobot
136	5	23.8	309	1	ACT3_ECHGR	Q03342	echinococcu	209	5	23.8	371	1	Y140_SYNY3	P73408	synechocyst
137	5	23.8	311	1	NOT1_RHILV	P08720	rhizobium l	210	5	23.8	372	1	ACT_LGMUR	P91754	lumbricus r
138	5	23.8	312	1	RT02_ACACA	P46753	acanthamoeb	211	5	23.8	373	1	ACT_CHOCR	P53499	chondrus cr
139	5	23.8	314	1	NOD1_RHILP	P23718	rhizobium l	212	5	23.8	374	1	ACTM_STRPU	P12431	strongyloce
140	5	23.8	320	1	Y678_HETJA	Q56091	methanococ	213	5	23.8	375	1	ACT1_ACACA	P02578	acanthamoeb
141	5	23.8	322	1	ACT_PROCL	P45521	procambarus	214	5	23.8	375	1	ACT1_DICDI	P02577	dictyosteli
142	5	23.8	323	1	CTTG_HAEDU	Q30827	haemophilus	215	5	23.8	375	1	ACT1_ECHGR	P35432	echinococcu
143	5	23.8	324	1	GCI_MOUSE	P01868	mus musculus	216	5	23.8	375	1	ACT1_FUGRU	P35484	fugu rubrip
144	5	23.8	327	1	ACT3_ARTSX	P18602	artemia sp.	217	5	23.8	375	1	ACT1_MAIZE	P27131	naegleria f
145	5	23.8	328	1	CAHB_HUMAN	O75493	homo sapien	218	5	23.8	375	1	ACT1_NAEFO	P12715	oxytricha n
146	5	23.8	328	1	CAHB_SHEEP	Q95203	ovis aries	219	5	23.8	375	1	ACT1_OXINO	P33468	oxytricha n
147	5	23.8	328	1	YCDU_ECGLI	P75910	escherichia	220	5	23.8	375	1	ACT1_OXYTR	P43239	pneumocysti
148	5	23.8	330	1	GCAA_MOUSE	P01863	mus musculus	221	5	23.8	375	1	ACT1_SCHCO	Q9Y702	schizophyll
149	5	23.8	331	1	ACT_COSCS	P30161	costaria co	222	5	23.8	375	1	ACT1_SCHCO	O9Y701	suillus bov
150	5	23.8	332	1	ACT1_SOLTU	P33587	solanum tub	223	5	23.8	375	1	ACT1_SUIBO	P10992	tetrahymena
151	5	23.8	333	1	TAL2_YEAST	P53228	saccharomyc	224	5	23.8	375	1	ACT1_TETTH	P07827	dictyosteli
152	5	23.8	333	1	Y237_SYNY3	Q55480	synechocyst	225	5	23.8	375	1	ACT2_DICDI	P53485	fugu rubrip
153	5	23.8	335	1	GCRB_MOUSE	P01864	mus musculus	226	5	23.8	375	1	ACT2_FUGRU	P55805	oxytricha n
154	5	23.8	335	1	GVPN_ANAEF	P55150	anabaena fl	227	5	23.8	375	1	ACT2_OXINO	P33469	oxytricha n
155	5	23.8	336	1	ACT1_LYCPS	Q96482	lycopersico	228	5	23.8	375	1	ACT2_SCHCO	Q9Y896	schizophyll
156	5	23.8	336	1	ACT2_LYCPS	Q96483	lycopersico	229	5	23.8	375	1	ACT2_SCHCO	Q9Y896	schizophyll
157	5	23.8	336	1	ACT2_SOLTU	P93586	solanum tub	230	5	23.8	375	1	ACT2_SUIBO	P07829	dictyosteli
158	5	23.8	336	1	ACT2_TOBAC	P93374	nicotiana t	231	5	23.8	375	1	ACT2_DICDI	P07829	dictyosteli
159	5	23.8	336	1	ACT3_LYCPS	Q96484	lycopersico	232	5	23.8	375	1	ACT3_FUGRU	P53486	fugu rubrip
160	5	23.8	336	1	ACT4_LYCPS	Q96484	lycopersico	233	5	23.8	375	1	ACT8_DICDI	P07830	dictyosteli
161	5	23.8	336	1	ACT4_TOBAC	P93372	nicotiana t	234	5	23.8	375	1	ACTA_PHYPO	P02576	physarum po
162	5	23.8	336	1	ACT5_SOLTU	P81228	solanum tub	235	5	23.8	375	1	ACTB_CRIGR	P48975	crisetulus
163	5	23.8	336	1	ACT5_TOBAC	P93371	nicotiana t	236	5	23.8	375	1	ACTB_CYPCA	P12714	cyprinus ca
164	5	23.8	336	1	ACT6_TOBAC	P93376	nicotiana t	237	5	23.8	375	1	ACTB_CYPCA	P02570	homo sapien
165	5	23.8	336	1	ACT7_TOBAC	P93375	nicotiana t	238	5	23.8	375	1	ACTB_RABIT	P29751	oryctolagus
166	5	23.8	336	1	ACT8_SOLTU	P81229	solanum tub	239	5	23.8	375	1	ACTB_SALSA	O42161	saimeo salar
167	5	23.8	336	1	ACT9_SOLTU	P93584	solanum tub	240	5	23.8	375	1	ACTC_BRABE	Q93129	branchiosto
168	5	23.8	336	1	AGLF_RHIME	Q92316	rhizobium m	241	5	23.8	375	1	ACTC_BRAFL	Q93131	branchiosto
169	5	23.8	336	1	GCB_MOUSE	P01866	mus musculus	242	5	23.8	375	1	ACTC_BRALA	O17503	branchiosto
170	5	23.8	337	1	A85A_MYCUL	P58248	mycobacteri	243	5	23.8	375	1	ACTC_STYPL	Q00215	styela plic
171	5	23.8	337	1	ACT4_SOLTU	P93585	solanum tub	244	5	23.8	375	1	ACTG_CEPAC	Q9uvw9	cephalospor
172	5	23.8	337	1	YJMC_BACSU	O34736	bacillus su	245	5	23.8	375	1	ACTG_EMENI	P20359	emerella
173	5	23.8	338	1	FLAG_PSEAE	P51464	pseudomonas	246	5	23.8	375	1	ACTG_HUMAN	P02571	homo sapien
174	5	23.8	338	1	YBOL_YEAST	P38200	saccharomyc	247	5	23.8	375	1	ACTG_PENCH	Q9urs0	penicillium
175	5	23.8	340	1	TPHS_SULAC	P17118	sulfolobus	248	5	23.8	375	1	ACT_AJCEA	P53455	ajellomyces
176	5	23.8	341	1	VP3_GFLV	P17768	grapevine f	249	5	23.8	375	1	ACT_BOYCI	O13419	botrytis ci
177	5	23.8	342	1	COBT_METSQ	Q9x7f4	methyllobact	250	5	23.8	375	1	ACT_COPCI	Q9uvx4	cryptospori
178	5	23.8	342	1	ILVC_BACSU	P37253	bacillus su	251	5	23.8	375	1	ACT_CRYNE	P48465	cryptococcu
179	5	23.8	343	1	BBUR_BORBR	O06703	bordetella	252	5	23.8	375	1	ACT_FUCDI	P53502	fucus disti

253	1	ACT_NEUCR	375	5	23.8	375	P78711 neutrospora	326	5	23.8	376	1	ACT_BROMA	376	1	P90689 brugia mala
254	1	ACT_PUARI	375	5	23.8	375	P53689 phaffia rho	327	5	23.8	376	1	ACT_CANAL	376	1	P14235 candida alb
255	1	ACT_PUGR	375	5	23.8	375	P50138 puccinia gr	328	5	23.8	376	1	ACT_CANDU	376	1	Q9uv28 candida dub
256	5	ACT_SCHPO	375	5	23.8	375	P10989 schizosacch	329	5	23.8	376	1	ACT_CERGI	376	1	O17320 crassostrea
257	1	ACT_THELA	375	5	23.8	375	P10365 thermomyces	330	5	23.8	376	1	ACT_CRYPV	376	1	P26183 cryptospori
258	5	ACT1_AEDAE	376	5	23.8	376	P49128 aedes aegypt	331	5	23.8	376	1	ACT_ENTHI	376	1	P11426 entamoeba h
259	5	ACT1_ARTSX	376	5	23.8	376	P18600 artemia sp.	332	5	23.8	376	1	ACT_FUCVE	376	1	Q39758 lucius vesic
260	1	ACT1_BOMMO	376	5	23.8	376	P07836 bombyx mori	333	5	23.8	376	1	ACT_HYDAT	376	1	P17126 hydra atten
261	5	ACT1_CAEEL	376	5	23.8	376	P10987 caenorhabdi	334	5	23.8	376	1	ACT_MANSE	376	1	P49871 manduca sex
262	1	ACT1_DROME	376	5	23.8	376	P10987 drosophila	335	5	23.8	376	1	ACT_MAYDE	376	1	O16808 mayetiola d
263	1	ACT1_HELER	376	5	23.8	376	P53462 heliocidari	336	5	23.8	376	1	ACT_PICAN	376	1	O74258 pichia angu
264	5	ACT1_LUMTE	376	5	23.8	376	P92182 lumbricus t	337	5	23.8	376	1	ACT_PLAMG	376	1	Q26065 placopecten
265	5	ACT1_LYPTI	376	5	23.8	376	P53465 lytechinus	338	5	23.8	376	1	ACT_TAESO	376	1	P14227 taenia soli
266	5	ACT1_ONCVO	376	5	23.8	376	P30162 onchocerca	339	5	23.8	376	1	ACT_TOXGO	376	1	P53476 toxoplasma
267	1	ACT1_PEA	376	5	23.8	376	P30164 pisum sativ	340	5	23.8	376	1	R51C_HUMAN	376	1	O43502 homo sapien
268	5	ACT1_PHYIN	376	5	23.8	376	P22131 phytophthor	341	5	23.8	376	1	XYNA_BACOV	376	1	P49942 bacteroides
269	5	ACT1_PLAFA	376	5	23.8	376	P10988 plasmodium	342	5	23.8	377	1	ACT1_ORYLA	377	1	O98972 oryzias lat
270	1	ACT1_PODCA	376	5	23.8	376	P11112 podocoryne	343	5	23.8	377	1	ACT1_ORYSA	377	1	P13362 oryza sativ
271	5	ACT1_SACKO	376	5	23.8	376	O18499 saccoglossu	344	5	23.8	377	1	ACT1_SORBI	377	1	P33504 sorghum bic
272	5	ACT1_SCHMA	376	5	23.8	376	P53470 schistosoma	345	5	23.8	377	1	ACT1_SOYBN	377	1	P02581 glycine max
273	5	ACT1_STREN	376	5	23.8	376	P10990 strongyloce	346	5	23.8	377	1	ACT1_TOBAC	377	1	Q05214 nicotiana t
274	5	ACT2_ARTSX	376	5	23.8	376	P18601 artemia sp.	347	5	23.8	377	1	ACT1_XENLA	377	1	P04751 xenopus lae
275	5	ACT2_RACDO	376	5	23.8	376	P45885 bactrocera	348	5	23.8	377	1	ACT1_XENGL	377	1	P26197 absidia gla
276	5	ACT2_BOMMO	376	5	23.8	376	P07837 bombyx mori	349	5	23.8	377	1	ACT2_XENLA	377	1	P53492 arabidopsis
277	5	ACT2_CAEEL	376	5	23.8	376	P10984 caenorhabdi	350	5	23.8	377	1	ACT2_XENLA	377	1	P10995 xenopus lae
278	5	ACT2_DIPDE	376	5	23.8	377	P53456 diphyllobot	351	5	23.8	377	1	ACT2_XENTR	377	1	P20399 xenopus tro
279	5	ACT2_DROME	376	5	23.8	377	P02572 drosophila	352	5	23.8	377	1	ACT3_ARATH	377	1	P53493 arabidopsis
280	5	ACT2_SCHGR	376	5	23.8	377	Q03341 echinococu	353	5	23.8	377	1	ACT3_DIPDE	377	1	P53457 diphyllobot
281	5	ACT2_LUMTE	376	5	23.8	377	P92176 lumbricus t	354	5	23.8	377	1	ACT3_PEA	377	1	P46258 pisum sativ
282	5	ACT2_LYPTI	376	5	23.8	377	P53466 lytechinus	355	5	23.8	377	1	ACT3_SOLTU	377	1	P30167 solanum tub
283	5	ACT2_ONCVO	376	5	23.8	377	P30163 onchocerca	356	5	23.8	377	1	ACT3_XENLA	377	1	P04752 xenopus lae
284	5	ACT2_PEA	376	5	23.8	377	P30165 pisum sativ	357	5	23.8	377	1	ACT4_XENLA	377	1	P53494 arabidopsis
285	5	ACT2_PLAFA	376	5	23.8	377	P14883 plasmodium	358	5	23.8	377	1	ACT6_SOLTU	377	1	P30168 solanum tub
286	5	ACT2_SACKO	376	5	23.8	377	O18500 saccoglossu	359	5	23.8	377	1	ACT7_SOLTU	377	1	P30169 solanum tub
287	5	ACT2_SCHMA	376	5	23.8	377	P53471 schistosoma	360	5	23.8	377	1	ACTA_CHICK	377	1	P08023 gallus gall
288	5	ACT2_TETPY	376	5	23.8	377	P10993 tetrahymena	361	5	23.8	377	1	ACTA_HUMAN	377	1	P03996 homo sapien
289	5	ACT3_BACDO	376	5	23.8	377	P45886 bactrocera	362	5	23.8	377	1	ACTB_ARATH	377	1	P53496 arabidopsis
290	5	ACT3_BOMMO	376	5	23.8	377	P04829 bombyx mori	363	5	23.8	377	1	ACTB_SOLTU	377	1	P30171 solanum tub
291	5	ACT3_DROME	376	5	23.8	377	P53501 drosophila	364	5	23.8	377	1	ACTC_ARATH	377	1	P53497 arabidopsis
292	5	ACT3_HELAM	376	5	23.8	377	Q25010 helicoverpa	365	5	23.8	377	1	ACTC_FUGRU	377	1	P53480 fugu rubrip
293	5	ACT3_LIMPO	376	5	23.8	377	P41340 limulus pol	366	5	23.8	377	1	ACTC_HUMAN	377	1	P04270 homo sapien
294	5	ACT3_ORYSA	376	5	23.8	377	P17299 oryza sativ	367	5	23.8	377	1	ACTD_SOLTU	377	1	P30173 solanum tub
295	5	ACT3_PODCA	376	5	23.8	377	P41113 podocoryne	368	5	23.8	377	1	ACTS_CARAU	377	1	P49055 carassius a
296	5	ACT3_SOYBN	376	5	23.8	377	P02580 glycine max	369	5	23.8	377	1	ACTS_CYPCA	377	1	P33479 cyprinus ca
297	5	ACT4_ARTSX	376	5	23.8	377	P18603 artemia sp.	370	5	23.8	377	1	ACTS_FUGRU	377	1	P53481 fugu rubrip
298	5	ACT4_BOMMO	376	5	23.8	377	Q27250 bombyx mori	371	5	23.8	377	1	ACTS_HUMAN	377	1	P02568 homo sapien
299	5	ACT4_CAEEL	376	5	23.8	377	P10986 caenorhabdi	372	5	23.8	377	1	ACTT_FUGRU	377	1	P53482 fugu rubrip
300	5	ACT4_DROME	376	5	23.8	377	P02574 drosophila	373	5	23.8	377	1	ACTI_CHLRE	377	1	P53498 chlamydomon
301	5	ACT5_BACDO	376	5	23.8	377	P45887 bactrocera	374	5	23.8	377	1	ACTI_COLSC	377	1	O65315 coleochaete
302	5	ACT5_CHICK	376	5	23.8	377	P53478 gallus gall	375	5	23.8	377	1	ACTI_CYAME	377	1	O65315 cyanidiosch
303	5	ACT5_DROME	376	5	23.8	377	P10981 drosophila	376	5	23.8	377	1	ACTI_MESVI	377	1	O85316 mesostigma
304	5	ACT5_XENLA	376	5	23.8	377	P33505 xenopus lae	377	5	23.8	377	1	ACTI_VOLCA	377	1	P20904 volvox cart
305	5	ACT6_DROME	376	5	23.8	377	P02575 drosophila	378	5	23.8	377	1	ACTI_HALRO	377	1	P53460 halocynthia
306	5	ACT7_ORYSA	376	5	23.8	379	P17300 oryza sativ	379	5	23.8	378	1	ACT2_HALRO	378	1	P27130 halocynthia
307	5	ACT8_XENLA	376	5	23.8	379	P53506 xenopus lae	380	5	23.8	378	1	ACT2_MOLOC	378	1	Q25472 molgula oca
308	5	ACT8_LIMPO	376	5	23.8	379	P41339 limulus pol	381	5	23.8	378	1	ACTM_BRAFL	378	1	Q93132 branchiost
309	5	ACT8_STRPU	376	5	23.8	379	P53472 strongyloce	382	5	23.8	378	1	ACTM_CIOSA	378	1	O15998 ciona savig
310	5	ACTB_STRPU	376	5	23.8	379	P53473 strongyloce	383	5	23.8	378	1	ACTM_MOLOC	378	1	P53467 molgula oca
311	5	ACTC_HALRO	376	5	23.8	379	P53461 halocynthia	384	5	23.8	378	1	ACTM_STYCL	378	1	P36198 styela clav
312	5	ACTC_PISOC	376	5	23.8	379	P12716 pisaster oc	385	5	23.8	378	1	ACTN_STYCL	378	1	P53475 styela clav
313	5	ACTD_STRPU	376	5	23.8	379	Q07903 strongyloce	386	5	23.8	379	1	ACTI_SCHDU	379	1	O65314 scherffelia
314	5	ACTE_STRPU	376	5	23.8	379	P10991 strongyloce	387	5	23.8	379	1	ACT2_ORYSA	379	1	P17298 oryza sativ
315	5	ACTF_STRPU	376	5	23.8	379	P53474 strongyloce	388	5	23.8	379	1	ACTM_BRABE	379	1	Q93130 branchiost
316	5	ACTG_STRPU	376	5	23.8	379	P18499 strongyloce	389	5	23.8	379	1	ACTM_STYPL	379	1	Q00214 styela plic
317	5	ACTH_HUMAN	376	5	23.8	379	P12718 homo sapien	390	5	23.8	380	1	ACTI_DAUC	380	1	P23343 daucus caro
318	5	ACTI_APLCA	376	5	23.8	379	P17304 aplysia cal	391	5	23.8	380	1	ACTM_BRALA	380	1	O17502 branchiost
319	5	ACTM_HELER	376	5	23.8	379	P53463 heliocidari	392	5	23.8	380	1	METB_HELPJ	380	1	Q92mw7 heliocobact
320	5	ACTM_HELTP	376	5	23.8	379	P53464 heliocidari	393	5	23.8	380	1	METB_HELPJ	380	1	P56069 heliocobact
321	5	ACTM_PISOC	376	5	23.8	379	P12717 pisaster oc	394	5	23.8	380	1	PLVA_COLGL	380	1	Q00374 colletotric
322	5	ACTX_FUGRU	376	5	23.8	379	P53483 fugu rubrip	395	5	23.8	381	1	ACT2_DAUCA	381	1	P23344 daucus caro
323	5	ACTY_LIMPO	376	5	23.8	379	P41341 limulus pol	396	5	23.8	385	1	ACTZ_PNECA	385	1	P42023 pneumocysti
324	5	ACT_ACHBI	376	5	23.8	379	P26182 achlya bise	397	5	23.8	387	1	ASPP_AEDAE	387	1	Q03168 aedes aegypt
325	5	ACT_BIOGL	376	5	23.8	379	P92179 blomphalar	398	5	23.8	389	1	DNAJ_METWA	389	1	P55515 methanosarc

399	5	23.8	390	1	CATD_BOVIN	P80209	bos taurus	472	5	23.8	467	1	CBPA_DICDI	P35085	dictyosteli
400	5	23.8	390	1	RPP_MUMPI	P19717	mumps virus	473	5	23.8	467	1	VNS5_TSNVL	P26003	tomato spot
401	5	23.8	390	1	KRI_RHISN	P55642	rhizobium s	474	5	23.8	468	1	T10A_HUMAN	O00220	homo sapien
402	5	23.8	391	1	GALL_MOUSE	Q90000	mus musculus	475	5	23.8	471	1	IF2G_MOUSE	P41091	homo sapien
403	5	23.8	391	1	RPP_MUMPM	P16072	mumps virus	476	5	23.8	471	1	IF2G_MOUSE	Q92001	mus musculus
404	5	23.8	392	1	RPP_MUMPM	P16595	mumps virus	477	5	23.8	471	1	IF2H_MOUSE	Q92002	mus musculus
405	5	23.8	392	1	PAF2_BOVIN	P79106	bos taurus	478	5	23.8	473	1	CB1R_POEBO	P56371	poephila gu
406	5	23.8	392	1	PAF2_HUMAN	Q99487	homo sapien	479	5	23.8	474	1	GATA_SCHPO	Q13837	schizosacch
407	5	23.8	392	1	SC11_HUMAN	Q96015	homo sapien	480	5	23.8	474	1	PRTP_MOUSE	P16675	mus musculus
408	5	23.8	393	1	GC1M_MOUSE	P01869	mus musculus	481	5	23.8	475	1	IF2G_DROME	Q24208	drosophila
409	5	23.8	394	1	CIW3_HUMAN	O14649	homo sapien	482	5	23.8	479	1	OPN4_HUMAN	P50874	homo sapien
410	5	23.8	395	1	NEUA_STRAG	Q53598	streptococ	483	5	23.8	479	1	ORC5_YEAST	P08619	saccharomyc
411	5	23.8	396	1	BRB2_RAT	P25023	rattus norv	484	5	23.8	480	1	PRTP_HUMAN	O16619	homo sapien
412	5	23.8	396	1	CARP_NEUCR	Q01294	neurospora	485	5	23.8	480	1	Y523_XYLEA	O68639	xyliella fas
413	5	23.8	396	1	SOB_HAEIN	P44535	haemophilus	486	5	23.8	482	1	TISD_HUMAN	P47974	homo sapien
414	5	23.8	396	1	YD18_YEAST	Q12185	saccharomyc	487	5	23.8	482	1	TISD_HUMAN	P10577	rhizobium m
415	5	23.8	396	1	YWB2_BACSU	P39587	bacillus su	488	5	23.8	484	1	NTRC_RHINE	Q02920	glycine max
416	5	23.8	397	1	P622_YEAST	Q12471	saccharomyc	489	5	23.8	485	1	NO70_SOYBN	P39844	bacillus su
417	5	23.8	397	1	LHX3_HUMAN	Q90000	homo sapien	490	5	23.8	491	1	PBP_RACSU	P49318	ricinus com
418	5	23.8	398	1	APL1_HUMAN	Q14791	homo sapien	491	5	23.8	492	1	CAT2_RICCO	P14010	emericaella
419	5	23.8	398	1	TRA5_RHINE	O52873	rhizobium m	492	5	23.8	498	1	GATA_EMENI	Q03413	xenopus lae
420	5	23.8	399	1	GCAM_MOUSE	P01865	mus musculus	493	5	23.8	498	1	YDHF_SCHPO	Q23360	schizosacch
421	5	23.8	400	1	PAT1_MUSVI	P50449	mustela vis	494	5	23.8	501	1	LOH2_MERTH	O26978	methanobact
422	5	23.8	400	1	V063_FOWPV	P21971	foalpoxy vir	495	5	23.8	505	1	MALO_SYNAV	P72785	synecocyst
423	5	23.8	402	1	PAT1_BOVIN	P13909	bos taurus	496	5	23.8	506	1	Y157_HUMAN	P51786	homo sapien
424	5	23.8	402	1	PAT1_HUMAN	P05121	bos taurus	497	5	23.8	507	1	IRX3_MOUSE	P81067	mus musculus
425	5	23.8	402	1	PAT1_PIG	P79335	sus scrofa	498	5	23.8	508	1	MKK1_YEAST	P32490	saccharomyc
426	5	23.8	402	1	PKR_CHLPN	Q92705	chlamydia p	499	5	23.8	508	1	VU1_HPVIA	P03099	human papil
427	5	23.8	405	1	GBM_MOUSE	P01867	mus musculus	500	5	23.8	509	1	FLAA_HELPU	Q03843	helicobacte
428	5	23.8	407	1	CATD_RAT	P24268	rattus norv	501	5	23.8	510	1	PRO1_LISMO	P23224	listeria mo
429	5	23.8	408	1	RUN3_MOUSE	P03389	rauscher sp	502	5	23.8	510	1	PRO2_LISMO	P34025	listeria mo
430	5	23.8	409	1	ENV_RSFFV	Q64131	m runt-rela	503	5	23.8	512	1	AMV1_DEBOC	P19269	debaromyce
431	5	23.8	410	1	CATD_MYCCE	P18242	mus musculus	504	5	23.8	512	1	FLAA_HELPU	Q90000	helicobacte
432	5	23.8	411	1	Y360_MYCCE	Q49426	mycoplasma	505	5	23.8	514	1	TSSP_HUMAN	Q90000	homo sapien
433	5	23.8	412	1	CATD_HUMAN	P07339	homo sapien	506	5	23.8	517	1	DN23_HUMAN	P08619	homo sapien
434	5	23.8	414	1	PRVC_THEAC	Q90000	thermoplas	507	5	23.8	518	1	RNA5_HUMAN	Q90000	homo sapien
435	5	23.8	415	1	RUN3_HUMAN	Q13761	h runt-rela	508	5	23.8	521	1	VENV_DHV1	P27427	dhori virus
436	5	23.8	416	1	ENV3_CAEEL	Q90000	caenorhabdi	509	5	23.8	522	1	GES1_YEAST	P41913	saccharomyc
437	5	23.8	417	1	TRAL_HUMAN	Q13077	homo sapien	510	5	23.8	523	1	REF3_LACIA	Q90000	lactococcus
438	5	23.8	417	1	EXFL_LACIA	Q90000	lactococcus	511	5	23.8	524	1	LCYE_ARATH	Q38932	arabidopsis
439	5	23.8	418	1	EXFL_SHIFL	P37781	shigella fl	512	5	23.8	526	1	CYB3_DROME	Q90000	drosophila
440	5	23.8	419	1	HAT1_HUMAN	Q14929	homo sapien	513	5	23.8	536	1	YRN3_CAEEL	Q09606	caenorhabdi
441	5	23.8	419	1	NIFS_FRASE	Q925X5	frankia sp	514	5	23.8	541	1	PELP_YERPS	P11278	versinia ps
442	5	23.8	421	1	EXOF_RHIME	Q02728	rhizobium m	515	5	23.8	543	1	APPA_BACSU	P42061	bacillus su
443	5	23.8	421	1	HMHD_AERPE	Q90000	aeropyrum p	516	5	23.8	545	1	YB90_YEAST	P38346	saccharomyc
444	5	23.8	423	1	SVY_ECOLI	P00951	escherichia	517	5	23.8	553	1	SECY_MAIZE	O63066	zea mays (m
445	5	23.8	427	1	IM44_SCHPO	O60084	schizosacch	518	5	23.8	556	1	INX7_CAEEL	Q21123	caenorhabdi
446	5	23.8	428	1	YFJD_ECOLI	P37908	escherichia	519	5	23.8	564	1	HEMA_IAPCZ	P19696	influenza a
447	5	23.8	432	1	ENV2_MOUSE	P11370	mus musculus	520	5	23.8	564	1	HEMA_IAPCZ	P19696	influenza a
448	5	23.8	432	1	PNL1_MOUSE	P35821	mus musculus	521	5	23.8	568	1	PELP_ERCA	P14005	erinia car
449	5	23.8	432	1	PNL1_RAT	P20417	rattus norv	522	5	23.8	571	1	PRIN_DEINA	Q90000	deinococcus
450	5	23.8	432	1	YFLL_HAEIN	P44717	haemophilus	523	5	23.8	578	1	YFLL_HAEIN	P44038	haemophilus
451	5	23.8	434	1	CPSS_STRCP	P39041	streptomyce	524	5	23.8	581	1	YG25_ARCFU	O28648	archaeoglob
452	5	23.8	435	1	MAIC_STRCP	Q04698	streptococ	525	5	23.8	581	1	YW04_SCHPO	Q90000	schizosacch
453	5	23.8	435	1	PNL1_HUMAN	P18031	homo sapien	526	5	23.8	586	1	CYDD_HAEIN	P45082	haemophilus
454	5	23.8	436	1	EPFL_RALSO	P58594	raistonia s	527	5	23.8	586	1	KPKY_BACPY	P51182	bacillus ps
455	5	23.8	436	1	EPFL_RALSO	Q95411	raistonia s	528	5	23.8	586	1	RRPO_BWYVF	Q90000	beet wester
456	5	23.8	437	1	ADFP_HUMAN	Q95411	homo sapien	529	5	23.8	587	1	CVA3_RHIME	Q92300	rhizobium m
457	5	23.8	440	1	BCHN_HELMO	Q90000	heliobacilli	530	5	23.8	589	1	Y567_HAEIN	P45221	haemophilus
458	5	23.8	442	1	SNA_ECOLI	Q46812	escherichia	531	5	23.8	590	1	PEPF_BORBU	O51264	borrelia bu
459	5	23.8	443	1	FLVS_FLAME	Q47899	flavobacter	532	5	23.8	593	1	SLT_HAEIN	P44888	haemophilus
460	5	23.8	443	1	Y082_SYNAV	Q55803	synecocyst	533	5	23.8	593	1	HM21_HUMAN	Q90000	homo sapien
461	5	23.8	444	1	CADB_ECOLI	P23891	escherichia	534	5	23.8	599	1	LAM2_HUMAN	Q03252	homo sapien
462	5	23.8	449	1	PKR1_PYRHO	O57979	pyrococcus	535	5	23.8	599	1	Y0U3_CAEEL	P30639	caenorhabdi
463	5	23.8	450	1	ADFP_BOVIN	Q90000	bos taurus	536	5	23.8	608	1	GHR_CHICK	Q02092	gallus gall
464	5	23.8	451	1	TBG_ENTHI	P54401	entamoeba h	537	5	23.8	610	1	RHO_METVA	P45835	mycobacteri
465	5	23.8	451	1	VU10_HSVJ7	P52524	human herpe	538	5	23.8	614	1	RPB1_METVA	P41557	methanococ
466	5	23.8	454	1	VNUC_THOGV	P89216	thogoto vir	539	5	23.8	614	1	YQZ2_CAEEL	Q09320	caenorhabdi
467	5	23.8	457	1	IRF7_MOUSE	P70434	mus musculus	540	5	23.8	617	1	Y41_BOVIN	Q90000	bos taurus
468	5	23.8	461	1	SL1_YEAST	Q04673	saccharomyc	541	5	23.8	618	1	VEL_HPV63	Q90000	human papil
469	5	23.8	463	1	VL2_HPV13	Q02275	human papil	542	5	23.8	621	1	PLAK_MOUSE	Q07847	mus musculus
470	5	23.8	464	1	VNS5_TSNVL	P26002	tomato spot	543	5	23.8	625	1	AMVG_NEUCR	Q02257	mus musculus
471	5	23.8	467	1	ATPX_BACFI	P25075	bacillus fi	544	5	23.8	636	1	ENV_MCFE	P15073	mink cell f

545	5	23.8	636	1	RPBL_METJA	Q60181 methanococ	618	5	23.8	879	1	E4L1_MOUSE	Q922h5 mus musculu
546	5	23.8	638	1	YD14_SCHPO	Q10237 schizosacch	619	5	23.8	879	1	YDBH_ECOLI	P52645 escherichia
547	5	23.8	640	1	ENV_RMGFP	P06445 raucher mi	620	5	23.8	881	1	E4L1_HUMAN	Q9h490 homo sapien
548	5	23.8	641	1	SYB2_THEMA	Q9wzj9 thermotoga	621	5	23.8	888	1	M3KC_MOUSE	O67000 mus musculu
549	5	23.8	641	1	FIB2_PETMA	P33573 petromyzo	622	5	23.8	888	1	M3KC_RAT	O63796 rattus norv
550	5	23.8	641	1	IMD_ARTGO	Q44052 arthrobacte	623	5	23.8	894	1	RRN6_YEAST	P32781 saccharomyc
551	5	23.8	647	1	ARS_CHLRE	P14217 chlamydomon	624	5	23.8	899	1	VP3_EHDV1	P27281 epizootic h
552	5	23.8	653	1	PPEI_HUMAN	O44829 homo sapien	625	5	23.8	899	1	VP3_EHDV1	P33474 epizootic h
553	5	23.8	656	1	YC26_PORPU	P51392 porphyra pu	626	5	23.8	901	1	VP3_BTIV1	P12435 bluetongue
554	5	23.8	658	1	CPT2_HUMAN	P23786 homo sapien	627	5	23.8	901	1	VP3_BTIV1	O65749 bluetongue
555	5	23.8	658	1	CPT2_MOUSE	P52825 mus musculu	628	5	23.8	901	1	VP3_BTIV1	O65750 bluetongue
556	5	23.8	658	1	CPT2_RAT	P18886 rattus norv	629	5	23.8	901	1	VP3_BTIV1	P20608 bluetongue
557	5	23.8	661	1	ENV_MLVCB	P08360 cas-br-e mu	630	5	23.8	901	1	VP3_BTIV1	P35359 bluetongue
558	5	23.8	662	1	Y4ID_RHISN	P55487 rhizobium s	631	5	23.8	901	1	VP3_BTIV1	P56582 bluetongue
559	5	23.8	665	1	ENV_MLVHO	P03385 moloney mur	632	5	23.8	901	1	VP3_BTIV1	O65748 bluetongue
560	5	23.8	666	1	ENV_MLVHO	P21436 homulv muri	633	5	23.8	906	1	R1R1_HCMVA	P16782 human cytom
561	5	23.8	672	1	YCDR_ECOLI	P75906 escherichia	634	5	23.8	908	1	DPO1_BORBU	O51498 borrelia bu
562	5	23.8	675	1	ENV_MLVF5	P03390 friend muri	635	5	23.8	911	1	DLG1_RAT	O62696 rattus norv
563	5	23.8	675	1	NED1_MOUSE	P33215 mus musculu	636	5	23.8	916	1	SECA_TREPA	O83394 treponema p
564	5	23.8	676	1	ENV_MLVFP	P26803 friend muri	637	5	23.8	917	1	SLAP_THETH	P32917 saccharomyc
565	5	23.8	681	1	TRF_MANSE	P22297 manduca sex	638	5	23.8	926	1	STES_YEAST	P29074 homo sapien
566	5	23.8	682	1	PRC_SALTY	P43669 salmonella	639	5	23.8	926	1	PTN4_HUMAN	Q92398 chlamydia p
567	5	23.8	682	1	SNK_MOUSE	P53351 mus musculu	640	5	23.8	926	1	PMP9_CHLPN	Q9wv92 mus musculu
568	5	23.8	682	1	SNK_RAT	Q9r012 rattus norv	641	5	23.8	929	1	E4L3_MOUSE	O92393 chlamydia p
569	5	23.8	685	1	SNK_HUMAN	Q9nyv3 homo sapien	642	5	23.8	930	1	PMP8_CHLPN	Q16512 homo sapien
570	5	23.8	685	1	SNWA_DICDI	P54705 dictyosteli	643	5	23.8	942	1	PKL1_HUMAN	P39523 saccharomyc
571	5	23.8	686	1	ENMD_HUMAN	O95936 homo sapien	644	5	23.8	943	1	PKL1_YEAST	O63433 rattus norv
572	5	23.8	686	1	KLC_STRPU	Q05090 strongyloce	645	5	23.8	946	1	PKL1_RAT	O51777 borrelia bu
573	5	23.8	689	1	DNL2_RICPR	Q926k9 rickettsia	646	5	23.8	950	1	UVRA_BORBU	O00339 homo sapien
574	5	23.8	690	1	VG42_BPMU	Q926k9 rickettsia	647	5	23.8	956	1	MTN2_HUMAN	O08746 mus musculu
575	5	23.8	700	1	YDHV_ECOLI	Q926k9 rickettsia	648	5	23.8	956	1	MTN2_MOUSE	P324618 saccharomyc
576	5	23.8	704	1	YBY1_YEAST	P76192 escherichia	649	5	23.8	956	1	YEF3_YEAST	O07075 homo sapien
577	5	23.8	711	1	FRE2_YEAST	P36033 saccharomyc	650	5	23.8	957	1	AMPE_HUMAN	O11171 mycobacteri
578	5	23.8	712	1	CYAB_BORPE	P18770 bordetella	651	5	23.8	968	1	MML2_MYCTU	Q92398 chlamydia p
579	5	23.8	715	1	AT12_HSV1F	P08314 herpes simp	652	5	23.8	975	1	CSE1_DROME	O92393 chlamydia p
580	5	23.8	718	1	MUS2_STAAU	P10230 herpes simp	653	5	23.8	990	1	YAA2_SCHPO	Q09796 schizosacch
581	5	23.8	719	1	EF2_METBU	Q926h5 staphylococ	654	5	23.8	991	1	SCA4_RICSI	Q9a177 rickettsia
582	5	23.8	730	1	EF2_METBU	Q93632 methanococ	655	5	23.8	992	1	EVC_HUMAN	P57679 homo sapien
583	5	23.8	730	1	EF2_METWT	Q93637 methanococ	656	5	23.8	1005	1	Y456_CHLTR	O84462 chlamydia t
584	5	23.8	735	1	CNG1_CHICK	Q90805 gallus gall	657	5	23.8	1010	1	SCA4_RICPA	O9a175 rickettsia
585	5	23.8	735	1	MCW5_XENLA	P58662 xenopus lae	658	5	23.8	1011	1	SCA4_RICAP	O9a183 rickettsia
586	5	23.8	738	1	PLAK_XENLA	P30998 xenopus lae	659	5	23.8	1011	1	SCA4_RICMN	O9a182 rickettsia
587	5	23.8	742	1	RMG5_YEAST	Q03656 saccharomyc	660	5	23.8	1012	1	SCA4_RICSL	O9a180 rickettsia
588	5	23.8	743	1	PLAK_HUMAN	P14923 homo sapien	661	5	23.8	1012	1	UBAL_SCHPO	O94609 schizosacch
589	5	23.8	757	1	MUS2_THEMA	Q9x105 thermotoga	662	5	23.8	1013	1	SCA4_RICRH	O9a181 rickettsia
590	5	23.8	758	1	PHSA_SALTY	P37600 salmonella	663	5	23.8	1018	1	SCA4_RICJA	Q9a179 rickettsia
591	5	23.8	768	1	DAB2_RAT	O88797 rattus norv	664	5	23.8	1018	1	YK2_CABEL	P41993 caenorhabdi
592	5	23.8	773	1	PAC2_PSES3	P15558 pseudomonas	665	5	23.8	1022	1	SCA4_RICCN	Q52658 rickettsia
593	5	23.8	774	1	STF_LAMB2	P03764 bacterioph	666	5	23.8	1026	1	PTPL_CABEL	P28191 caenorhabdi
594	5	23.8	780	1	MUS2_BORBU	O51125 borrelia bu	667	5	23.8	1067	1	SGG_DROME	P18431 drosophila
595	5	23.8	785	1	MUS2_BACHD	Q9x8a0 bacillus ha	668	5	23.8	1087	1	FAL3_HUMAN	O9y232 homo sapien
596	5	23.8	785	1	MUS2_BACSU	P94345 bacillus su	669	5	23.8	1093	1	PIAK_DICDI	P54577 dictyosteli
597	5	23.8	785	1	YHY2_YEAST	P38870 saccharomyc	670	5	23.8	1094	1	EMBA_MYCTU	P72060 mycobacteri
598	5	23.8	788	1	UGS3_SOITU	Q43847 solanum tub	671	5	23.8	1122	1	EX5C_ECOLI	P07648 escherichia
599	5	23.8	790	1	SYFB_CHILU	Q9pj18 chlamydia m	672	5	23.8	1122	1	RPOB_THECE	P31814 thermococcu
600	5	23.8	799	1	RSQ4_MOUSE	Q92268 mus musculu	673	5	23.8	1139	1	NGAP_HUMAN	Q9u1f2 homo sapien
601	5	23.8	801	1	41_XENLA	P11434 xenopus lae	674	5	23.8	1141	1	SRE2_HUMAN	Q12772 homo sapien
602	5	23.8	802	1	PEFC_SALTY	P37868 salmonella	675	5	23.8	1146	1	CCAS_RAT	Q02485 rattus norv
603	5	23.8	815	1	FAED_ECOLI	P06970 escherichia	676	5	23.8	1160	1	TFC3_YEAST	P34111 saccharomyc
604	5	23.8	815	1	AKH_HAETN	P44505 haemophilus	677	5	23.8	1195	1	RPOB_THEAC	Q03587 thermoplasm
605	5	23.8	816	1	AKH_BUCAL	P57290 buchnera ap	678	5	23.8	1199	1	NIEJ_ANASP	O06879 anabaena sp
606	5	23.8	819	1	AKIH_SERMA	P27725 serratia ma	679	5	23.8	1227	1	CLBE_BACTO	O85805 bacillus th
607	5	23.8	820	1	AKIH_ECOLI	P00561 escherichia	680	5	23.8	1253	1	MIS9_DROME	Q01989 drosophila
608	5	23.8	824	1	JIP2_HUMAN	Q13387 homo sapien	681	5	23.8	1262	1	GNRP_MOUSE	P27671 mus musculu
609	5	23.8	829	1	CAD3_HUMAN	P22223 homo sapien	682	5	23.8	1267	1	VL3_REOVJ	P17378 reovirus (t
610	5	23.8	830	1	JIP2_MOUSE	Q9ere9 mus musculu	683	5	23.8	1267	1	VL3_REOVJ	P17377 reovirus (t
611	5	23.8	835	1	OBP_VZVD	P09299 varicella-z	684	5	23.8	1267	1	MDR1_CRIGR	P21448 cricetus
612	5	23.8	849	1	DEGX_CABEL	Q09274 caenorhabdi	685	5	23.8	1276	1	MDR1_MOUSE	O67995 mus musculu
613	5	23.8	851	1	OBP_HSV11	P10193 herpes simp	686	5	23.8	1276	1	MDR2_CRIGR	P21449 cricetus
614	5	23.8	858	1	41_MOUSE	Q48193 mus musculu	687	5	23.8	1276	1	MDR1_HUMAN	O08183 homo sapien
615	5	23.8	859	1	M3KC_HUMAN	Q12852 homo sapien	688	5	23.8	1280	1	RPOC_MYCGE	P47582 mycoplasma
616	5	23.8	864	1	41_HUMAN	P11171 homo sapien	689	5	23.8	1292	1	CAIE_HUMAN	P39059 homo sapien
617	5	23.8	875	1	FOP1_YEAST	P41812 saccharomyc	690	5	23.8	1388	1		

691	5	23.8	1395	1	SP41_YEAST	P38904	saccharomyc	764	4	19.0	38	1	CPRP_CANPG	P81033	cancer pagu
692	5	23.8	1398	1	DNA2_SCHPO	O9uru2	schizosacch	765	4	19.0	39	1	PSAG_PEA	P20120	pisum sativ
693	5	23.8	1403	1	BIRE_MOUSE	O9r016	mus musculus	766	4	19.0	40	1	PURA_BACSU	O00829	bacillus su
694	5	23.8	1407	1	ROC_ECOLI	P00577	escherichia	767	4	19.0	41	1	RYK2_LYCES	P36493	lycopersico
695	5	23.8	1407	1	ROC_SALTY	O61197	salmonella	768	4	19.0	42	1	HM81_SHEEP	P14159	ovis aries
696	5	23.8	1434	1	PCCL_MOUSE	O61115	mus musculus	769	4	19.0	43	1	RL40_METHH	O26653	methanobact
697	5	23.8	1490	1	CKRT_HUMAN	O9nyv4	homo sapien	770	4	19.0	44	1	Y234_HAEIN	P43968	haemophilus
698	5	23.8	1522	1	PST1_SCHPO	O99750	schizosacch	771	4	19.0	45	1	LHA7_RHOAC	P35093	rhodopseu
699	5	23.8	1557	1	LM11_CAEEL	O18923	caenorhabdi	772	4	19.0	46	1	RL40_ARCFU	O28842	archaeoglob
700	5	23.8	1581	1	PRPB_HUMAN	O18923	homo sapien	773	4	19.0	47	1	RL40_ARCFU	P42354	archaeoglob
701	5	23.8	1607	1	LMGL_MOUSE	P02468	mus musculus	774	4	19.0	48	1	RM32_TOBAC	P12198	nicotiana t
702	5	23.8	1609	1	LMGL_HUMAN	P11047	homo sapien	775	4	19.0	49	1	RM32_TOBAC	P28804	spinacia ol
703	5	23.8	1687	1	CCAM_MUSDO	O25452	musca domes	776	4	19.0	50	1	VSH_WUMPA	P28085	mumps virus
704	5	23.8	1698	1	Y076_HUMAN	O14999	homo sapien	777	4	19.0	51	1	VSH_WUMPA	P28085	mumps virus
705	5	23.8	1699	1	POLN_LORDV	P54634	lordsdale v	778	4	19.0	52	1	RCSB_KLEAE	P05339	klebsiella
706	5	23.8	1770	1	R115_YEAST	P43565	saccharomyc	779	4	19.0	53	1	L503_MOUSE	O9wvdb7	mus musculus
707	5	23.8	1786	1	YCF1_ARATH	P56785	arabidopsis	780	4	19.0	54	1	L503_MOUSE	O9wvdb7	mus musculus
708	5	23.8	1818	1	Y294_HUMAN	O94822	homo sapien	781	4	19.0	55	1	Y4DN_RHISN	P55413	rhizobium s
709	5	23.8	1828	1	CUTL_SCHPO	P18296	schizosacch	782	4	19.0	56	1	CE3D_HYPCU	P50720	hyphantria
710	5	23.8	1856	1	GBF1_CRIGR	O9rid7	cricetulus	783	4	19.0	57	1	CE3E_HYPCU	P50721	hyphantria
711	5	23.8	1859	1	GBF1_HUMAN	O92538	homo sapien	784	4	19.0	58	1	CE3F_HYPCU	P50722	hyphantria
712	5	23.8	1873	1	CCAS_HUMAN	O13698	homo sapien	785	4	19.0	59	1	CE3G_HYPCU	P50723	hyphantria
713	5	23.8	1873	1	CCAS_RABIT	P07293	oryctolagus	786	4	19.0	60	1	RL32_RICPR	O9zch0	rickettsia
714	5	23.8	1901	1	YCF1_TOBAC	P12222	nicotiana t	787	4	19.0	61	1	VRPI_BBP7	P03704	bacterioph
715	5	23.8	1951	1	CJN3_HUMAN	O9n746	homo sapien	788	4	19.0	62	1	AMP1_MELGA	P80391	meleagris g
716	5	23.8	1951	1	CJN3_RAT	P08104	rattus norv	789	4	19.0	63	1	GULL_CHICK	P46156	gallus gall
717	5	23.8	1959	1	AGRI_RAT	P25304	rattus norv	790	4	19.0	64	1	YC33_CYAPA	P48373	cyanophora
718	5	23.8	1971	1	MC3A_MOUSE	O9wvdb7	mus musculus	791	4	19.0	65	1	YVCE_VACCC	P21123	vaccinia vi
719	5	23.8	2124	1	Y192_HUMAN	O93074	homo sapien	792	4	19.0	66	1	EMX1_MOUSE	O04742	mus musculus
720	5	23.8	2145	1	CYAA_PODAN	O01513	podospira a	793	4	19.0	67	1	RFC4_BPPI	P11189	bacterioph
721	5	23.8	2151	1	RRPL_HANTV	P23456	hantaan vir	794	4	19.0	68	1	Y07A_ARCFU	P58023	archaeoglob
722	5	23.8	2151	1	RRPL_SEOUB	P27314	seoul virus	795	4	19.0	69	1	ATPL_STRMU	P59783	streptococ
723	5	23.8	2156	1	RRPL_PUMOH	P27176	pumalia vir	796	4	19.0	70	1	QTX2_RAT	O64201	rattus norv
724	5	23.8	2223	1	CCAE_DISOM	P56699	discopyge o	797	4	19.0	71	1	PYRH_THETH	P43891	thermus aqu
725	5	23.8	2248	1	Y539_HUMAN	O60287	homo sapien	798	4	19.0	72	1	DMS4_PACDA	O93454	pachymedusa
726	5	23.8	2254	1	CCAG_RAT	O54898	rattus norv	799	4	19.0	73	1	FBS_MOUSE	O60791	mus musculus
727	5	23.8	2255	1	RRPL_SV5	O88434	simian viru	800	4	19.0	74	1	RL29_RICCN	O92484	rickettsia
728	5	23.8	2255	1	RRPL_SV5WR	O03996	simian viru	801	4	19.0	75	1	RL29_RICCN	O92484	rickettsia
729	5	23.8	2261	1	RRPL_MUPWM	P03929	mumps virus	802	4	19.0	76	1	CMO6_CONGE	P01522	conus geogr
730	5	23.8	2262	1	RRPL_P1ZHT	P26676	human para	803	4	19.0	77	1	DMS2_AGAN	O93222	agalactem
731	5	23.8	2269	1	RRPL_SV41	P53341	simian viru	804	4	19.0	78	1	Y055_NPVAC	P41459	autographa
732	5	23.8	2273	1	ABCR_HUMAN	P78363	homo sapien	805	4	19.0	79	1	Y889_ARCFU	O28584	archaeoglob
733	5	23.8	2441	1	CHP_MOUSE	P45481	mus musculus	806	4	19.0	80	1	BB11_SCHCO	P78742	schizophyl
734	5	23.8	2516	1	CCAD_DROME	O24270	drosophila	807	4	19.0	81	1	Y4DJ_RHISN	P54079	rhizobium s
735	5	23.8	2517	1	NCR2_HUMAN	O9y618	h nuclear r	808	4	19.0	82	1	Y8K6_BPP22	P14107	bacterioph
736	5	23.8	2555	1	PS3_BACSU	P39847	bacillus su	809	4	19.0	83	1	ATPB_STRDP	P21933	streptococ
737	5	23.8	2871	1	DESP_HUMAN	P15924	homo sapien	810	4	19.0	84	1	R31B_STRPN	O979c0	streptococ
738	5	23.8	3110	1	HD_RAT	P51111	rattus norv	811	4	19.0	85	1	Y072_BACAN	O9rmw2	bacillus an
739	5	23.8	3119	1	HD_MOUSE	P42859	mus musculus	812	4	19.0	86	1	Y897_RICPR	O9rmw2	bacillus an
740	5	23.8	3122	1	POZ_MOUSE	O61493	mus musculus	813	4	19.0	87	1	RSS_ARATH	P51427	arabidopsis
741	5	23.8	3125	1	POLG_PPVNA	P17766	p genome po	814	4	19.0	88	1	TATA_MYCTU	Q10703	mycobacteri
742	5	23.8	3140	1	POLG_PPVRA	P17767	p genome po	815	4	19.0	89	1	YQGD_ECOLI	P46879	escherichia
743	5	23.8	3140	1	POLG_PPVSK	O84934	p genome po	816	4	19.0	90	1	PLS_HAIIA	P82595	hallois la
744	5	23.8	3141	1	POLG_PPVD	P13529	p genome po	817	4	19.0	91	1	YVFD_VACCC	P20562	vaccinia vi
745	5	23.8	3144	1	HD_HUMAN	P42858	homo sapien	818	4	19.0	92	1	CYS2_PASHA	P72205	pasteurella
746	5	23.8	3148	1	HD_FUGRU	P51112	fuqu rubrip	819	4	19.0	93	1	R31B_STRPY	O9a016	streptococ
747	5	23.8	3163	1	POLG_TUMVQ	O02597	t genome po	820	4	19.0	94	1	AGA2_YEAST	P32781	saccharomyc
748	5	23.8	3164	1	POLG_TUMVJ	P89509	t genome po	821	4	19.0	95	1	IM08_YEAST	P57744	saccharomyc
749	5	23.8	3674	1	SPRC_HUMAN	O9nrc6	homo sapien	822	4	19.0	96	1	R35A_PYRAB	O9v1p2	pyrococcus
750	5	23.8	4092	1	DHCC_YEAST	P36022	saccharomyc	823	4	19.0	97	1	R35A_PYRHO	O74099	pyrococcus
751	5	23.8	4196	1	FAT1_SCHPO	O13490	schizosacch	824	4	19.0	98	1	R35A_PYRHO	P20299	pyrococcus
752	5	23.8	5147	1	FAT1_DROME	P33450	drosophila	825	4	19.0	99	1	AMSD_RAT	Q13106	homo sapien
753	5	23.8	5217	1	HTSL_COCCA	O01886	cochliobolu	826	4	19.0	100	1	AMSD_RAT	Q02356	rattus norv
754	5	23.8	5327	1	ACF7_HUMAN	O04920	homo sapien	827	4	19.0	101	1	RS15_ECOLI	P02371	escherichia
755	5	23.8	5430	1	ACF7_MOUSE	O9upn3	homo sapien	828	4	19.0	102	1	RS15_HAEIN	P44389	haemophilus
756	5	23.8	5430	1	ACF7_MOUSE	O9upn3	homo sapien	829	4	19.0	103	1	RS15_HAEIN	P41120	photobactu
757	4	19.0	12	1	SO15_BACSU	P80863	callinectes	830	4	19.0	104	1	TATA_MYCLE	P54079	mycobacteri
758	4	19.0	19	1	MDH_SHEPU	P81682	staphylococ	831	4	19.0	105	1	Y113_METEX	O05113	methylobact
759	4	19.0	30	1	OTCX_STAEP	P24244	escherichia	832	4	19.0	106	1	YDBJ_ECOLI	P52646	escherichia
760	4	19.0	30	1	YCB_ECOLI	O9mus3	mesostigma	833	4	19.0	107	1	CYC6_MONBR	O09099	monoraphidi
761	4	19.0	33	1	YCB_ECOLI	O9mus3	mesostigma	834	4	19.0	108	1	CYC6_MONBR	P57736	scenedesmus
762	4	19.0	37	1	LPYV_SERNA	P19937	serratia ma	835	4	19.0	109	1	DYLL_ANTCR	O02414	anthracidari
763	4	19.0	37	1	POLN_WERY	P13896	western equ	836	4	19.0	110	1	DYLL_CABEL	Q22799	caenorhabdi

837	89	1	DYL1_DROME	Q24117 drosophila	910	109	1	AQP2_TALEU	07740 talpa europ
838	89	1	DYL1_HUMAN	Q15701 homo sapien	911	109	1	COX3_ARBLI	Q33752 arabacia lix
839	89	1	DYL2_DROME	Q96860 drosophila	912	109	1	IATO_CABEL	Q4441 caenorhabdi
840	89	1	FL10_AQAE	Q67774 aquifex aeo	913	109	1	YOS1_BPP2	Q06426 bacteriophia
841	89	1	DBH_PASMU	Q08394 pasteurella	914	110	1	AGF1_RAT	Q35460 rattus norv
842	90	1	LITL_PHYSA	Q08948 phyllomedus	915	110	1	LV1C_MOUSE	P01725 mus musculu
843	90	1	LITP_PHYSA	Q08947 phyllomedus	916	110	1	RN_BACCI	P35078 bacillus ci
844	90	1	LPFA_HUMAN	Q09568 homo sapien	917	111	1	CO7R_MOUSE	Q61387 mus musculu
845	91	1	DYL1_CHLRE	Q39580 chlamydomon	918	111	1	LV2H_HUMAN	P01709 homo sapien
846	91	1	VC15_VACCC	P21099 vaccinia vi	919	111	1	LV2H_HUMAN	P01711 homo sapien
847	92	1	COR_BPPH8	P17651 bacteriophia	920	111	1	Y101_HUMAN	Q15004 homo sapien
848	92	1	HRK_MOUSE	P17658 mus musculu	921	111	1	Y847_TREPA	Q83819 treponema p
849	92	1	Y6D8_YERPE	P16161 yersinia pe	922	111	1	YC39_AQAE	Q67284 aquifex aeo
850	92	1	YR13_CORAM	Q45825 corynebacte	923	112	1	ANFG_AZUAE	Q68946 azomonas ma
851	93	1	HIS2_MYCLE	Q49786 mycobacteri	924	112	1	FLJE_AGRU5	Q44337 agrobacteri
852	93	1	HIS2_MYCTU	Q33257 mycobacteri	925	112	1	SO06_BOVIN	P80231 bos taurus
853	93	1	RLA0_PIG	Q29214 sus scrofa	926	113	1	FRY2_HUMAN	Q75474 homo sapien
854	94	1	RT19_PETHY	P27527 petunia hyb	927	113	1	LV1_CHICK	P04210 gallus gall
855	95	1	MINE_DEIRA	Q97508 deinococcus	928	113	1	MEAL_PIG	Q95313 sus scrofa
856	95	1	VE10_VACCC	P21050 vaccinia vi	929	113	1	NPFF_HUMAN	O15130 homo sapien
857	95	1	VE10_VACCV	P23373 vaccinia vi	930	113	1	RK20_NEPOL	Q97kx4 nephroselm
858	95	1	VE10_VARY	P33821 variola vir	931	113	1	RSBV_STROCO	Q9wv88 streptomyce
859	95	1	Y614_MYCLE	Q49760 mycobacteri	932	113	1	VL3_BP4	P08343 bovine papi
860	96	1	C553_HELPY	Q25229 helicobacte	933	113	1	YA53_HAETN	Q57498 haemophilus
861	96	1	C553_HELPY	Q25825 helicobacte	934	113	1	YWDK_BACSU	P29619 bacillus su
862	96	1	HOLI_BPA18	Q37975 bacteriophia	935	114	1	ALP2_RHOCA	P50923 rhodobacter
863	96	1	HOLI_BPA50	Q37977 bacteriophia	936	114	1	KEDA_ACTSL	P41249 actinomycec
864	96	1	NULM_WYXGL	Q924x0 myxine gluc	937	114	1	MIFH_BRUMA	P31850 brugia mala
865	96	1	Y87A_HAETN	Q86223 haemophilus	938	114	1	RSBV_WUCBA	Q4786 wuchereria
866	96	1	YNBC_ECOLI	P76144 escherichia	939	114	1	RSBV_LISMO	O85016 listeria mo
867	97	1	APC2_MOUSE	Q05020 mus musculu	940	114	1	SZ06_HUMAN	P80162 homo sapien
868	97	1	NULM_MOUSE	P03903 mus musculu	941	114	1	VG40_BPT4	P17171 bacteriophia
869	97	1	NULM_STRPU	P15554 strongyloce	942	114	1	YGI3_BACTU	P10024 bacillus th
870	98	1	FER2_NOSMU	P00249 nostoc musc	943	115	1	COX3_ARTSA	Q33845 artemia sal
871	98	1	NULM_SCYCA	O79409 scyllorhinu	944	115	1	IFIA_PYRAB	Q9v138 pyrococcus
872	98	1	NULM_SQUAC	Q92246 squalus aca	945	115	1	Y4HO_RHISN	P50359 rhizobium s
873	98	1	SL13_BOVIN	P79342 bos taurus	946	115	1	YACC_ECOLI	P23838 escherichia
874	98	1	SL13_MOUSE	P97352 mus musculu	947	115	1	YIDD_MYCPA	Q917m0 mycobacteri
875	99	1	ARG2_PHAU	P32292 phaseolus a	948	116	1	BR31_BRAPE	Q90435 brachydanio
876	99	1	NOS3_SHEEP	P79209 ovis aries	949	116	1	BR3A_CHICK	P55968 gallus gall
877	100	1	CMW1_SINY3	P72759 synchocyst	950	116	1	E411_ADEI2	P36708 human adeno
878	100	1	NSR2_PSEAE	Q01709 pseudomonas	951	116	1	IFIA_PYRHO	O59280 pyrococcus
879	100	1	RL14_PEA	P05638 pisum sativ	952	116	1	LVIA_MOUSE	P01723 mus musculu
880	101	1	PCHE_PSEAE	Q15107 pseudomonas	953	116	1	RL17_SYNPF	O24711 synchocococ
881	101	1	RNPO_PLEOS	P81762 pleurotonas	954	116	1	SP21_BACME	P35147 bacillus me
882	102	1	ACYM_CHICK	P07031 gallus gall	955	116	1	SL21_BACST	O32726 bacillus st
883	102	1	RK23_ODOSI	P49559 odontella s	956	116	1	ULC9_HCMVA	P16838 human cytom
884	102	1	TXOA_HADIN	Q9b1v9 hadronyche	957	116	1	WN53_EPTST	P28121 eptatretus
885	102	1	Y12K_MSVN	P14974 maize strea	958	117	1	EBRB_BACSU	O31791 bacillus su
886	103	1	ANPB_BOVIN	P13204 bos taurus	959	117	1	GHRL_MOUSE	Q9eqx0 mus musculu
887	103	1	NTPG_ENTHR	P44355 enterococcu	960	117	1	GHRL_RAT	Q9qv17 rattus norv
888	103	1	RL24_HAETN	P44352 haemophilus	961	117	1	LV0A_HUMAN	P04211 homo sapien
889	103	1	RL24_PASMU	Q9cl11 pasteurella	962	117	1	LV2A_MOUSE	P01728 mus musculu
890	103	1	TAT_BIV06	P19584 bovine immu	963	117	1	TAT_SIVS4	P12513 simian immu
891	103	1	TAT_BIV27	P19585 bovine immu	964	117	1	WNT6_STRPU	P28097 strongyloce
892	103	1	Y01A_BPT4	P39419 bacteriophia	965	117	1	YDHX_CITFR	P45516 citrobacter
893	104	1	YPC6_CLAPU	P22368 clavicaps p	966	118	1	RL19_ASTLO	P24772 astasia ion
894	104	1	Y376_MYCGE	P47616 mycoplasma	967	118	1	YC87_MYCPN	P75490 mycoplasma
895	105	1	VPR_HV2RO	P06938 human immu	968	119	1	IPL_BACSU	P29804 bacillus su
896	105	1	YBFE_STAAU	Q33602 staphylococ	969	119	1	YB82_YEAST	P38327 saccharomyc
897	105	1	YD15_HAETN	P71375 haemophilus	970	119	1	YCFE_ECOLI	P16950 escherichia
898	106	1	REV_HV1BN	P12485 human immu	971	119	1	YJ22_YEAST	P47003 saccharomyc
899	106	1	RL11_HAEDU	Q32613 haemophilus	972	120	1	VG19_BPMU	Q38646 bacteriophia
900	106	1	Y266_RICPR	P39dd9 rickettsia	973	120	1	YCV2_YEAST	P25363 saccharomyc
901	106	1	YAL7_YEAST	P39557 saccharomyc	974	120	1	YJ08_YEAST	P47106 saccharomyc
902	108	1	RS25_LYCES	P46301 lycopersico	975	121	1	LSM4_SCHFO	O14352 schizosacch
903	108	1	URE2_PROVU	P16123 proteus vul	976	121	1	INIT_HUMAN	P40305 homo sapien
904	109	1	AQP2_BOVIN	P79099 bos taurus	977	122	1	NTF2_ORYSA	Q9xj54 oryza sativ
905	109	1	AQP2_CANFA	P79144 canis fami	978	122	1	PA2N_CROSSA	P23559 crotalus sc
906	109	1	AQP2_DASNO	O77714 dasypus nov	979	122	1	VG20_BPMU2	O64214 mycobacteri
907	109	1	AQP2_DUGDU	O77714 dugong dugu	980	122	1	YMT9_YEAST	Q05226 mycobacteri
908	109	1	AQP2_HORSE	P79165 equus cabal	981	122	1	YMT9_YEAST	Q04502 saccharomyc
909	109	1	AQP2_RABIT	P79213 oryctolagus	982	123	1	CRCB_PYRAB	Q9v0x2 pyrococcus

```

883 4 19.0 123 1 CRCB_PYRHO
884 4 19.0 123 1 YML4_YEAST
885 4 19.0 124 1 PAND-AQAE
886 4 19.0 124 1 R35A-CABEL
887 4 19.0 124 1 RS12-BUCAI
888 4 19.0 124 1 UL39-HCMA
889 4 19.0 124 1 VB03-VACCC
890 4 19.0 124 1 YGDD-HAEN
891 4 19.0 124 1 Y027-BACAN
892 4 19.0 125 1 HV2D-HUMAN
893 4 19.0 125 1 RL7-HELPI
894 4 19.0 125 1 RL7-HELPI
895 4 19.0 125 1 SSAO-SALTY
896 4 19.0 125 1 YN25-CABEL
897 4 19.0 126 1 CYCP-RHOU
898 4 19.0 126 1 HV2A-HUMAN
899 4 19.0 126 1 YIGF-ECOLI
900 4 19.0 126 1 YIGF-SALTY

O59171 pyrococcus
P53926 saccharomyc
O66773 aquifex aeo
P49180 caenorhabdi
P57595 buchera ap
P16813 human cytom
P21000 vaccinia vi
P45019 haemophilus
Q9rn05 bacillus an
P01817 homo sapien
P56875 helicobacte
P55834 helicobacte
P74858 salmonella
P34584 caenorhabdi
P00144 rhodospiril
P01814 homo sapien
P27842 escherichia
P31139 salmonella

ALIGNMENTS
PRT; 230 AA.
FILL_YEAST STANDARD; PRT; 230 AA.
AC P38771;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE FILL protein, mitochondrial precursor.
GN FILL OR KIM4 OR YHR038W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MF8-1;
RX MEDLINE=98417448; PubMed=9746366;
RA Kanai T., Takeshita S., Atomi H., Unemura K., Ueda M., Tanaka A.;
RT "A regulatory factor, FILL, involved in derepression of the
RT isocitrate lyase gene in Saccharomyces cerevisiae -- a possible
RT mitochondrial protein necessary for protein synthesis in
RT mitochondria.";
RL Eur. J. Biochem. 256:212-220(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohldman P., Waterston R., Wilson R.,
RA Vaadin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994).
RN [3]
RP SIMILARITY TO RRF.
RX MEDLINE=96155066; PubMed=8563640;
RA Ouzounis C., Bork P., Casati G., Sander C.;
RT "New protein functions in yeast chromosome VIII.";
RL Protein Sci. 4:2424-2428(1995).
CC -!- FUNCTION: INVOLVED IN GLUCOSE REPRESSION/DEREPRESSION. MAY BE
CC NECESSARY FOR PROTEIN SYNTHESIS IN MITOCHONDRIA. MAY FUNCTION AS A
CC RIBOSOME RECYCLING FACTOR IN MITOCHONDRIA.
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- SIMILARITY: BELONGS TO THE RRF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
-----
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: AB016033; BAA31687.1; -
DR EMBL: U00062; AAB68906.1; -
DR PIR: S46737; S46737.
DR SGD: S0001080; FILL.
DR InterPro: IPR002661; RRF.
DR Pfam: PF01765; RRF; 1.
KW Protein biosynthesis; Transit peptide; Mitochondrion.
FT TRANSIT 1 24 MITOCHONDRION (POTENTIAL).
FT CHAIN 25 230 FILL PROTEIN.
SQ SEQUENCE 230 AA; 26406 MW; 9CBDC8C5F86F3008 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 RLFNRSF 15
DB 17 RLFNRSF 23

RESULT 2
ABFI_TRIRE
ID ABFI_TRIRE STANDARD; PRT; 500 AA.
AC Q92455;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Alpha-L-arabinofuranosidase precursor (EC 3.2.1.55) (Arabinosidase).
GN ABFI.
OS Trichoderma reesei (Hypocrea jecorina).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=51453;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=QM9414 / RUT C-30;
RX MEDLINE=96434473; PubMed=8837440;
RA Margolles-Clark E., Tenkanen M., Nakari-Setälä T., Penttilä M.;
RT "Cloning of genes encoding alpha-L-arabinofuranosidase and beta-
RT xylosidase from Trichoderma reesei by expression in Saccharomyces
RT cerevisiae.";
RL Appl. Environ. Microbiol. 62:3840-3846(1996).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing alpha-L-
CC arabinofuranoside residues in alpha-L-arabinosides.
CC -!- PATHWAY: INVOLVED IN DEGRADATION OF THE PLANT CELL WALL
CC POLYSACCHARIDE L-ARABINAN.
CC -!- SIMILARITY: BELONGS TO FAMILY 54 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
EMBL: Z69252; CAA93243.1; -
DR EMBL: Z69252; CAA93243.1; -
KW Hydrolase; Glycosidase; Glycoprotein; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 500 ALPHA-L-ARABINOFURANOSIDASE.
FT CARBOHYD 467 467 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 500 AA; 51115 MW; FC6B4DB03EE3C762 CRC64;

Query Match 33.3%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY 1 SGLSTF 7
DB 266 SGLSTF 272

RESULT 3
XYLI_TRIKO
ID XYLI_TRIKO STANDARD; PRT; 500 AA.
AC P48792;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Arabinofuranosidase/B-xylosidase precursor [includes: Alpha-L-
DE arabinofuranosidase (EC 3.2.1.55) (Arabinosidase); Beta-xylosidase
DE (EC 3.2.1.57) (1,4-beta-D-xylian xylohydrolase) (Xylan 1,4-beta-
DE xylosidase)].
GN XYLI.
OS Trichoderma koningii.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; Hypocrea.
OX NCBI_TaxID=55202;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=G-39;
RA Huang L.N., Hsue T.H., Lee Y.J.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal non-reducing alpha-L-
CC arabinofuranoside residues in alpha-L-arabinosides.
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-XYLIANS SO AS TO
CC REMOVE SUCCESSIVE D-XYLOSE RESIDUES FROM THE NON-REDUCING
CC TERMINI. IT ALSO HYDROLYSES XYLOBIOSE.
CC -!- SIMILARITY: BELONGS TO FAMILY 34 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U38561; AAA81024.1; -
KW Xylan degradation; Hydrolase; Glycosidase; Multifunctional enzyme;
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 500
FT CARBOHYD 467 467 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 500 AA; 51129 MW; 06DFC319AAFA1149 CRC64;

Query Match 33.3% Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred No. 4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGLSTF 7
DB 266 SGLSTF 272

RESULT 4
Y805_AQUAE
ID Y805_AQUAE STANDARD; PRT; 155 AA.
AC O66989;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_805.
GN AQ_805.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000708; AAC06954.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 155 AA; 18152 MW; 16933A2AC7344531 CRC64;

Query Match 28.6% Score 6; DB 1; Length 155;
Best Local Similarity 100.0%; Pred No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
DB 54 LFNRSF 59

RESULT 5
CCAC_CAVPO
ID CCAC_CAVPO STANDARD; PRT; 169 AA.
AC O35505;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1C subunit (Calcium
DE channel, L type, alpha-1 polypeptide, isoform 1, cardiac muscle
DE (fragment)).
GN CACNA1C OR CACNA1L OR CCHL1A1 OR CACH2 OR CACN2.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=uterus;
RA Collings P.L., Lundgren D.W., Kulp T.M., Shah P., Chang S.M.,
RA Chang A.S.;
RT "Gestational expression of voltage-dependent calcium channel subunits
RT in guinea pig uterus."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1C
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS. LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP),
CC PHENYLALKYLAMINES, BENZOTHAZEPINES, AND BY OMEGA-AGATOXIN-IIIA
CC (OMEGA-AGA-IIIA). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC CALCIUM CHANNELS CONTAINING THE ALPHA-1C SUBUNIT PLAY AN IMPORTANT
CC ROLE IN EXCITATION-CONTRACTION COUPLING IN THE HEART (BY
CC SIMILARITY).
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA

```

```

CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC
CC -!- DOMAIN: BINDING OF INTRACELLULAR CALCIUM THROUGH THE EF-HAND MOTIF
CC INHIBITS THE OPENING OF THE CHANNEL (BY SIMILARITY).
CC
CC -!- PTM: PHOSPHORYLATION BY CAPK ACTIVATES THE CHANNEL (BY
CC SIMILARITY).
CC
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNIT
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AF005938; AAB62890.1; -.
CC InterPro: IPR002077; Ca_channel.
CC InterPro: IPR000636; Cation_chan_non_lig.
CC InterPro: IPR001682; Channel_pore_Ca_Na.
CC Pfam: PF00520; Ion_trans; 1.
CC PRINTS: PR00167; CACHANNEL.
CC Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
CC Calcium channel; Glycoprotein; Repeat; Multigene family;
CC Calcium-binding; Phosphorylation.
CC NON_TER 1 1
CC TRANSMEM <1 11 S3 OF REPEAT IV (POTENTIAL).
CC DOMAIN 12 38 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 39 57 S4 OF REPEAT IV (POTENTIAL).
CC DOMAIN 58 76 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 77 96 S5 OF REPEAT IV (POTENTIAL).
CC DOMAIN 97 165 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 166 >169 S6 OF REPEAT IV (POTENTIAL).
CC BINDING <1 11 TO DIHYDROPYRIDINES (BY SIMILARITY).
CC BINDING 144 >169 TO DIHYDROPYRIDINES (BY SIMILARITY).
CC BINDING 158 >169 TO PHENYLALKYLAMINES (BY SIMILARITY).
CC SITE 130 130 CALCIUM ION SELECTIVITY AND PERMEABILITY
CC (BY SIMILARITY).
CC CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC NON_TER 169 169
CC SEQUENCE 169 AA; 19514 MW; 138E88E510D30CB8 CRC64;
CC
CC Query Match 28.6%; Score 6; DB 1; Length 169;
CC Best Local Similarity 100.0%; Pred. No. 17;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 6 TFFRLF 11
CC DB 40 TFFRLF 45
CC
CC RESULT 6
CC CCAC_CHICK STANDARD; PRT; 177 AA.
CC AC Q73707;
CC DT 15-JUL-1999 (Rel. 38, Created)
CC DT 15-JUL-1999 (Rel. 38, Last sequence update)
CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
CC DE Voltage-dependent L-type calcium channel alpha-1C subunit (CHCACHA1C)
CC DE (Fragment).
CC GN CACNA1C.
CC OS Gallus gallus (Chicken).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
CC OC Gallus.
CC OX NCBI_TaxID=9031;

```

```

RN
RP SEQUENCE FROM N.A.
RC STRAIN-WHITE LEHORN; TISSUE-Basilar papilla;
RX MEDLINE=98070847; PubMed=9405708;
RA Kollmar R., Montgomery L.G., Fak J., Henry L.J., Hudspeth A.J.;
RT "Predominance of the alpha1d subunit in L-type voltage-gated Ca2+
RL channels of hair cells in the chicken's cochlea.";
Proc. Natl. Acad. Sci. U.S.A. 94:14883-14888(1997).
CC -!- FUNCTION: THE ISOFORM ALPHA-1C GIVES RISE TO L-TYPE CALCIUM
CC CURRENTS. LONG-LASTING (L-TYPE) CALCIUM CHANNELS BELONG TO THE
CC "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP. CALCIUM CHANNELS CONTAINING
CC THE ALPHA-1C SUBUNIT PLAY AN IMPORTANT ROLE IN EXCITATION-
CC CONTRACTION COUPLING IN THE HEART (BY SIMILARITY).
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- DOMAIN: BINDING OF INTRACELLULAR CALCIUM THROUGH THE EF-HAND MOTIF
CC INHIBITS THE OPENING OF THE CHANNEL (BY SIMILARITY).
CC -!- PTM: PHOSPHORYLATION BY CAPK ACTIVATES THE CHANNEL (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AF027610; AAC08311.1; -.
CC InterPro: IPR002077; Ca_channel.
CC InterPro: IPR002111; cat_chan_non_lig.
CC InterPro: IPR000636; Cation_chan_non_lig.
CC InterPro: IPR001682; Channel_pore_Ca_Na.
CC Pfam: PF00520; Ion_trans; 1.
CC PRINTS: PR00167; CACHANNEL.
CC Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
CC Calcium channel; Glycoprotein; Repeat; Multigene family;
CC Calcium-binding; Phosphorylation.
CC NON_TER 1 1
CC REPEAT <1 >177 IV.
CC TRANSMEM <1 13 S3 OF REPEAT IV (POTENTIAL).
CC DOMAIN 14 26 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 27 45 S4 OF REPEAT IV (POTENTIAL).
CC DOMAIN 46 64 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 65 84 S5 OF REPEAT IV (POTENTIAL).
CC DOMAIN 85 153 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 154 >177 S6 OF REPEAT IV (POTENTIAL).
CC BINDING 132 >177 TO DIHYDROPYRIDINES (BY SIMILARITY).
CC BINDING 146 >177 TO PHENYLALKYLAMINES (BY SIMILARITY).
CC SITE 118 118 CALCIUM ION SELECTIVITY AND PERMEABILITY
CC (BY SIMILARITY).
CC CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC NON_TER 177 177
CC SEQUENCE 177 AA; 19957 MW; 84CB4656D78AAF23 CRC64;
CC
CC Query Match 28.6%; Score 6; DB 1; Length 177;
CC Best Local Similarity 100.0%; Pred. No. 17;
CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

QY      6 TFFRLF 11
Db      28 TFFRLF 33

RESULT 7
Y340_AQUAE
ID Y540_AQUAE STANDARD; PRT; 214 AA.
AC O66819;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AQ_540.
GN AQ_540.
OS Aquifex aeolicus.
OC Bacteria; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus";
RL Nature 392:353-358(1998).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0056 (MARC) FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000694; AAC06780.1;
DR InterPro; IPR002771; UPF0056.
DR Pfam; PF01914; UPF0056; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 47 67 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 122 142 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 185 205 POTENTIAL.
SQ SEQUENCE 214 AA; 23572 MW; DDIA9947A5DE931D CRC64;

Query Match 28.6%; Score 6; DB 1; Length 214;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 SLSTFF 8
Db      171 SLSTFF 176

RESULT 8
CCAD_MOUSE
ID CCAD_MOUSE STANDARD; PRT; 281 AA.
AC Q99246;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Voltage-dependent L-type calcium channel alpha-1D subunit (Calcium
DE channel, L type, alpha-1 polypeptide isoform 2) (Fragment).
GN CACNA1D OR CACNA1A2 OR CCHLA2 OR CACH3 OR CACH4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;

```

```

[1]
RN SEQUENCE FROM N.A. (ISOFORM CACH3B).
RP STRAIN=ICR; TISSUE=Ovary.
RX MEDLINE=91056091; PubMed=2173707;
RA Perez-Reyes E., Wei X., Castellano A., Birnbaumer L.;
RT "Molecular diversity of L-type calcium channels. Evidence for
RT alternative splicing of the transcripts of three non-allelic genes.";
RL J. Biol. Chem. 265:20430-20436(1990).
CC -1- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1D
CC GIVES RISE TO L-TYPE CALCIUM CURRENTS, LONG-LASTING (L-TYPE)
CC CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA)
CC GROUP. THEY ARE BLOCKED BY DIHYDROPYRIDINES (DHP), AGATOXIN-IIIa
CC PHENALKYLAMINES, BENZOTHIAZEPINES, AND BY OMEGA-AGATOXIN-IIIa
CC (OMEGA-AGA-IIIa). THEY ARE HOWEVER INSENSITIVE TO OMEGA-CONOTOXIN-
CC GVIA (OMEGA-CTX-GVIA) AND OMEGA-AGATOXIN-IVA (OMEGA-AGA-IVA).
CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE FORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS ARE PRODUCED BY
CC ALTERNATIVE SPLICING (PROBABLE). THE SEQUENCE SHOWN HERE IS THAT
CC OF CACH3B.
CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS, AT EVERY THIRD POSITION.
CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M57975; AAA63292.1;
DR MGD; MGI:88293; Cactnald.
DR InterPro; IPR002077; Ca_channel.
DR InterPro; IPR000636; Cation_chan_non_lig.
DR InterPro; IPR001682; Channel_pore_Ca_Na.
DR Pfam; PF00520; Ion_trans_1.
DR PRINTS; PR00167; CACHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Alternative splicing.
FT NON_TER 1 1
FT REPEAT 45 >281
FT DOMAIN 2 58
FT TRANSMEM 59 77
FT DOMAIN 78 92
FT TRANSMEM 93 112
FT DOMAIN 113 119
FT TRANSMEM 120 141
FT DOMAIN 142 151
FT TRANSMEM 152 171
FT DOMAIN 172 190
FT TRANSMEM 191 210
FT DOMAIN 211 277
FT TRANSMEM 278 >281
FT SITE 244 244
FT BINDING <1 18
FT TO DIHYDROPYRIDINES (BY SIMILARITY).

```

FT BINDING 258 >281 TO DIHYDROPYRIDINES (BY SIMILARITY).
 FT BINDING 270 >281 TO PHENYLALKYLAMINES (BY SIMILARITY).
 FT NON_TER 281
 SO SEQUENCE 281 AA; 32615 MW; 6F317P54D81E3F6P CRC64;
 Query Match 28.6%; Score 6; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 TFFRLP 11
 Db 154 TFFRLP 159
 RESULT 9
 ID CCAS_CHICK STANDARD; PRT; 281 AA.
 AC 042398;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Voltage-dependent L-type calcium channel alpha-1S subunit (Fragment).
 GN CACNA1S.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WHITE LSHORN; TISSUE=Intestinal epithelium;
 RA Xu J., Norman A.W., Henry H.L., de Boland A.R., Zanillo L.P.;
 RT "Molecular characterization of an L-type calcium channel in chick
 intestinal epithelial cells."
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THE ISOFORM ALPHA-1S GIVES RISE TO L-TYPE CALCIUM
 CURRENTS. LONG-LASTING (L-TYPE) CALCIUM CHANNELS BELONG TO THE
 "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP. THEY ARE BLOCKED BY
 DIHYDROPYRIDINES (DHP), PHENYLALKYLAMINES, BENZOTHAZEPINES, AND
 BY OMEGA-AGATOXIN-IIIA (OMEGA-AGA-IIIA). THEY ARE HOWEVER
 INSENSITIVE TO OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-GVIA) AND OMEGA-
 AGATOXIN-IVA (OMEGA-AGA-IVA). CALCIUM CHANNELS CONTAINING THE
 ALPHA-1S SUBUNIT PLAY AN IMPORTANT ROLE IN EXCITATION-CONTRACTION
 COUPLING IN SKELETAL MUSCLE (BY SIMILARITY).
 CC -1- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
 COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
 IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
 FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
 SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
 CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
 LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY. AN
 ADDITIONAL GAMMA SUBUNIT IS PRESENT ONLY IN SKELETAL MUSCLE L-TYPE
 CHANNEL (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
 HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
 POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
 PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
 SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
 CC -1- DOMAIN: THE LOOP BETWEEN REPEATS II AND III INTERACTS WITH THE
 RYANODINE RECEPTOR, AND IS THEREFORE IMPORTANT FOR CALCIUM RELEASE
 FROM THE ENDOPLASMIC RETICULUM NECESSARY FOR MUSCLE CONTRACTION
 (BY SIMILARITY).
 CC -1- PFM: PHOSPHORYLATION BY CAPK STIMULATES THE CALCIUM CHANNEL
 FUNCTION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNIT
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF007877; AAB63206.1; .
 DR InterPro; IPR002077; Ca_channel.
 DR InterPro; IPR002111; Cat_channel_TripL.
 DR InterPro; IPR000636; Cation_chan_non_lig.
 DR InterPro; IPR001682; Channel_pore_Ca_Na.
 DR Pfam; PF00520; ion_trans_1.
 DR PRINTS; PR00167; CACHANNEL.
 DR Ionic channel; transmembrane; Ion transport; Voltage-gated channel;
 KW Calcium channel; Glycoprotein; Repeat; Multigene family;
 KW Calcium-binding; Phosphorylation.
 FT NON_TER 1
 FT REPEAT <1 8 III.
 FT REPEAT 45 >281 IV.
 FT TRANSMEM <1 5 S6 OF REPEAT III (POTENTIAL).
 FT DOMAIN 6 58 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 59 77 S1 OF REPEAT IV (POTENTIAL).
 FT DOMAIN 78 92 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 93 112 S2 OF REPEAT IV (POTENTIAL).
 FT DOMAIN 113 120 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 121 139 S3 OF REPEAT IV (POTENTIAL).
 FT DOMAIN 140 152 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 153 171 S4 OF REPEAT IV (POTENTIAL).
 FT DOMAIN 172 190 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 191 210 S5 OF REPEAT IV (POTENTIAL).
 FT DOMAIN 211 277 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 278 >281 S6 OF REPEAT IV (POTENTIAL).
 FT BINDING <1 17 TO DIHYDROPYRIDINES (BY SIMILARITY).
 FT BINDING 258 >281 TO DIHYDROPYRIDINES (BY SIMILARITY).
 FT BINDING 270 >281 TO PHENYLALKYLAMINES (BY SIMILARITY).
 FT SITE 244 244 CALCIUM ION SELECTIVITY AND PERMEABILITY
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 281
 SO SEQUENCE 281 AA; 32729 MW; 46BE70FA44D9166B CRC64;
 Query Match 28.6%; Score 6; DB 1; Length 281;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 TFFRLP 11
 Db 154 TFFRLP 159
 RESULT 10
 ID SCF_CHICK STANDARD; PRT; 287 AA.
 AC 009108;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Kit ligand precursor (C-kit ligand) (Stem cell factor) (SCF) (Mast
 cell growth factor) (MGF).
 GN KITLG.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93273244; PubMed=7684722;
 RA Zhou J., Ontaki M., Sakurai M.;
 RT "Sequence of a cDNA encoding chicken stem cell factor."
 RL Gene 127:269-270(1993).
 CC -1- FUNCTION: STIMULATES THE PROLIFERATION OF MAST CELLS. ABLE TO
 AUGMENT THE PROLIFERATION OF BOTH MYELOID AND LYMPHOID
 HEMATOPOIETIC PROGENITORS IN BONE MARROW CULTURE. MEDIATES ALSO
 CELL-CELL ADHESION. ACTS SYNERGISTICALLY WITH OTHER CYTOKINES,
 PROBABLY INTERLEUKINS.

CC -!- SUBUNIT: HOMODIMER, NON-COVALENTLY LINKED (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. ALSO EXISTS AS A
 CC SECRETED SOLUBLE FORM.
 CC -!- DEVELOPMENTAL STAGE: ACTS IN THE EARLY STAGES OF HEMATOPOIESIS.
 CC -!- SIMILARITY: BELONGS TO THE SCF FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D13516; BAA02733.1; -;
 CC PIR; JN0637; JN0637.
 CC InterPro; IPR003452; SCF.
 CC Pfam; PF02404; SCF; 1.
 CC Growth factor; Glycoprotein; Transmembrane; Signal; Cell adhesion.
 CC SIGNAL 1 25
 CC CHAIN 26 287
 CC DOMAIN 26 225
 CC TRANSMEM 226 246
 CC DOMAIN 247 287
 CC DISULFID 29 117
 CC DISULFID 68 167
 CC CARBOHYD 100 100
 CC CARBOHYD 149 149
 CC CARBOHYD 200 200
 CC SEQUENCE 287 AA; 33328 MW; 6AE8556ADC152578 CRC64;
 CC -----
 CC Query Match 28.6%; Score 6; DB 1; Length 287;
 CC Best Local Similarity 100.0%; Pred. No. 27;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 7 FFRLFN 12
 CC |||||
 CC Db 144 FFRLFN 149
 CC -----
 CC RESULT 11
 CC CEBB_MOUSE STANDARD; PRT; 296 AA.
 CC AC P28033;
 CC DT 01-AUG-1992 (Rel. 23, Created)
 CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
 CC DE CCAAT/enhancer binding protein beta (C/EBP beta) (Interleukin-6-
 CC dependent binding protein) (IL-6DBP) (Liver-enriched transcriptional
 CC activator) (LAP) (AGP/EBP).
 CC CEBPB.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RC STRAIN=BAUB/C; TISSUE=Liver;
 CC RX MEDLINE=91061773; PubMed=1701020;
 CC RA Chang C.J., Chen T.T., Lei H.Y., Chen D.S., Lee S.C.;
 CC RT "Molecular cloning of a transcription factor, AGP/EBP, that belongs
 CC to members of the C/EBP family."
 CC RL Mol. Cell. Biol. 10:6642-6653(1990).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=91357470; PubMed=1840554;
 CC RX Cao Z., Umek R.M., McKnight S.L.;
 CC RT "Regulated expression of three C/EBP isoforms during adipose
 CC conversion of 3T3-L1 cells."
 CC RL Genes Dev. 5:1538-1552(1991).
 CC [3]
 CC RN SEQUENCE OF 1-11 FROM N.A.
 CC RP SEQUENCE=Liver;
 CC RN

EX MEDLINE=95322031; PubMed=7598808;
 RA Chang C.J., Shen B.J., Lee S.C.;
 RT "Autoregulated induction of the acute-phase response transcription
 RT factor gene, agp/ebp."
 RL Nucleic Acids Res. 14:529-537(1995).
 CC -!- FUNCTION: INVOLVED IN INTERLEUKIN-6 SIGNAL TRANSDUCTION,
 CC INCLUDING THE TRANSCRIPTIONAL ACTIVATION OF ACUTE-PHASE GENES.
 CC -!- SUBUNIT: BINDS DNA AS A DIMER AND CAN FORM STABLE HETERODIMERS
 CC WITH C/EBP ALPHA AND GAMMA.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: TO OTHER BZIP PROTEINS. STRONG, TO OTHER C/EBP
 CC PROTEINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X62600; CAA44484.1; -;
 CC EMBL; M61007; AAA37192.1; -;
 CC PIR; A37279; A37279.
 CC PIR; A36366; A36366.
 CC TRANSFAC; T00017; -;
 CC MGD; MGI:88373; Ceppb.
 CC InterPro; IPR001871; bZIP.
 CC Pfam; PF00170; bZIP; 1.
 CC SMART; SM00338; BRU2; 1.
 CC Transcription regulation; Activator; DNA-binding; Nuclear protein;
 CC Trans-acting factor.
 CC DOMAIN 120 129 PRO-RICH.
 CC DOMAIN 170 191 PRO/SER-RICH.
 CC DNA_BIND 228 244 BASIC MOTIF.
 CC DOMAIN 257 285 LEUCINE-ZIPPER.
 CC SEQUENCE 296 AA; 31445 MW; 827AC4AFC209A89 CRC64;
 CC -----
 CC Query Match 28.6%; Score 6; DB 1; Length 296;
 CC Best Local Similarity 100.0%; Pred. No. 27;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 SGLST 6
 CC |||||
 CC Db 174 SGLST 179
 CC -----
 CC RESULT 12
 CC CEBB_RAT STANDARD; PRT; 297 AA.
 CC AC P21272;
 CC DT 01-MAY-1991 (Rel. 18, Created)
 CC DT 01-MAY-1991 (Rel. 18, Last sequence update)
 CC DT 15-JUL-1998 (Rel. 36, Last annotation update)
 CC DE CCAAT/enhancer binding protein beta (C/EBP beta) (Interleukin-6-
 CC dependent binding protein) (IL-6DBP) (Liver-enriched transcriptional
 CC activator) (LAP) (Silencer factor B) (SF-B) (C/EBP-related protein 2).
 CC GN CEBPB OR SFB OR CRP2.
 CC OS Rattus norvegicus (Rat).
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC OX NCBI_TaxID=10116;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP MEDLINE=91029495; PubMed=2171780;
 CC RA Poli V., Mancini F.P., Cortese R.;
 CC RT "IL-6DBP, a nuclear protein involved in interleukin-6 signal
 CC transduction, defines a new family of leucine zipper proteins related
 CC to C/EBP."
 CC RL Cell 63:643-653(1990).
 CC [2]
 CC RN SEQUENCE FROM N.A.
 CC RP

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=90269225; PubMed=2112087;
 RA Akira S., Ishihara H., Sugita T., Tanabe O., Kinoshita S.,
 RA Nishio Y., Nakajima T., Hirano T., Kishimoto T.;
 RT "A nuclear factor for IL-6 expression (NF-IL6) is a member of a C/EBP
 RT family.";
 RL EMBO J. 9:1897-1906(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Copley V.E., Collier R.E., Connor R., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P., Dunn M.,
 RA Ellington A.G., Griffiths J.A., Fraser A., French L., Garner P.,
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levenson M.H., Leversha M., Lloyd K., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Shownkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -!- FUNCTION: IMPORTANT TRANSCRIPTIONAL ACTIVATOR IN THE REGULATION OF
 CC GENES INVOLVED IN IMMUNE AND INFLAMMATORY RESPONSES. SPECIFICALLY
 CC BINDS TO AN IL-1 RESPONSE ELEMENT IN THE IL-6 GENE. NF-IL6 ALSO
 CC BINDS TO REGULATORY REGIONS OF SEVERAL ACUTE-PHASE AND CYTOKINE
 CC GENES. IT PROBABLY PLAYS A ROLE IN THE REGULATION OF ACUTE-PHASE
 CC REACTION, INFLAMMATION AND HEMOPOIESIS. THE CONSENSUS RECOGNITION
 CC SITE IS 5'-T(T/C)GNGNAA(T/G)-3'.
 CC -!- SUBUNIT: BINDS DNA AS A DIMER AND CAN FORM STABLE HETERODIMERS
 CC WITH C/EBP ALPHA, DELTA AND GAMMA.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN THE LUNG, KIDNEY,
 CC AND SPLEEN.
 CC -!- SIMILARITY: TO OTHER BZIP PROTEINS. STRONG, TO OTHER C/EBP
 CC PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X52560; CAA36794.1; --
 CC DR EMBL; AL161937; CAC14276.1; --
 CC DR PIR; S12788; S12788.
 CC DR TRANSFAC; T00581; --
 CC DR MIM; 189965; --
 CC DR InterPro; IPR001871; BZIP.
 CC DR Pfam; PF00170; BZIP; 1.
 CC DR SMART; SM00338; BRLZ; 1.
 CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
 CC DOMAIN 162 170
 FT

FT DNA_BIND 277 293 BASIC MOTIF.
 FT DOMAIN 306 334 LEUCINE-ZIPPER.
 FT CONFLICT 253 G -> A (IN REF. 2).
 SQ SEQUENCE 345 AA; 36091 MW; C4D7A476CACC717D CRC64;
 Query Match 28.6%; Score 6; DB 1; Length 345;
 Best local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SGLSLT 6
 Db 221 SGLSLT 226
 RESULT 15
 CEBB_BOVIN
 ID CEBB_BOVIN STANDARD; PRT; 348 AA.
 AC Q02755;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE CCAAT/enhancer binding protein beta (C/EBP beta).
 GN CEBPB.
 OS Bos taurus (Bovinae).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese black;
 RX MEDLINE=97203913; PubMed=9051485;
 RA Yamaoka I., Taniguchi Y., Sasaki Y.;
 RT "Nucleotide sequence of bovine C/EBP beta gene.";
 RL J. Anim. Sci. 75:587-597(1997).
 CC -!- FUNCTION: INVOLVED IN INTERLEUKIN-6 SIGNAL TRANSDUCTION, INCLUDING
 CC THE TRANSCRIPTIONAL ACTIVATION OF ACUTE-PHASE GENES (BY
 CC SIMILARITY)
 CC -!- SUBUNIT: BINDS DNA AS A DIMER AND CAN FORM STABLE HETERODIMERS
 CC WITH C/EBP ALPHA AND GAMMA (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: TO OTHER BZIP PROTEINS. STRONG, TO OTHER C/EBP
 CC PROTEINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; D82985; BAA20096.1; --
 CC DR InterPro; IPR001871; BZIP.
 CC DR Pfam; PF00170; BZIP; 1.
 CC DR SMART; SM00338; BRLZ; 1.
 CC Transcription regulation; Activator; DNA-binding; Nuclear protein.
 CC DOMAIN 13 19 POLY-PRO.
 FT DOMAIN 162 170 POLY-PRO.
 FT DOMAIN 224 232 POLY-SER.
 FT DOMAIN 247 250 POLY-ALA.
 FT DNA_BIND 280 296 BASIC MOTIF.
 FT DOMAIN 309 337 LEUCINE-ZIPPER.
 SQ SEQUENCE 348 AA; 36390 MW; 7059D2F87B474CB4 CRC64;
 Query Match 28.6%; Score 6; DB 1; Length 348;
 Best local Similarity 100.0%; Pred. No. 32;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SGLSLT 6
 Db 222 SGLSLT 227

Wed Nov 13 14:19:10 2002

us-09-833-017b-4.oligo.rsp

Page 16

Search completed: November 5, 2002, 11:05:40
Job time : 20.6418 secs

GenCore version 5.1.3
Copyright (C) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2002, 10:58:27 ; Search time 13.791 Seconds
(without alignments)
263.424 Million cell updates/sec

Title: US-09-833-017B-4
Perfect score: 21
Sequence: 1 SGSLSTFFFLNRSFTQALGK 21

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

SPTREMBL_19:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	100.0	46	2	Q99Q15
2	21	100.0	46	2	Q9APK7
3	15	71.4	43	2	Q9APK6
4	7	33.3	253	13	Q90315
5	7	33.3	287	13	Q90314
6	7	33.3	929	5	Q9VR32
7	6	28.6	51	12	Q9IG33
8	6	28.6	69	12	Q41074
9	6	28.6	69	12	Q91PP4
10	6	28.6	122	16	Q9RU75
11	6	28.6	134	12	Q84484
12	6	28.6	140	4	Q9BSC0
13	6	28.6	177	15	Q9QRM3
14	6	28.6	180	16	Q98FX0
15	6	28.6	181	15	Q901F3
16	6	28.6	186	10	O04571

17	6	28.6	186	11	Q62146
18	6	28.6	191	11	Q9D077
19	6	28.6	200	5	O16314
20	6	28.6	200	6	Q95LH6
21	6	28.6	201	15	Q83184
22	6	28.6	202	16	Q9PLC4
23	6	28.6	204	11	Q9EPW1
24	6	28.6	205	5	O15772
25	6	28.6	205	5	O15777
26	6	28.6	213	16	O84471
27	6	28.6	246	5	Q18894
28	6	28.6	249	2	Q9RE46
29	6	28.6	271	16	Q9K8F9
30	6	28.6	285	16	Q9KDP6
31	6	28.6	287	16	Q97F64
32	6	28.6	288	16	Q9PB74
33	6	28.6	289	5	O01509
34	6	28.6	297	11	Q64106
35	6	28.6	300	2	Q93UN6
36	6	28.6	300	16	Q91259
37	6	28.6	314	1	P94910
38	6	28.6	318	4	Q9UM08
39	6	28.6	329	16	O83509
40	6	28.6	336	2	O87519
41	6	28.6	340	3	P78846
42	6	28.6	342	5	O9GRL0
43	6	28.6	345	4	O96IH2
44	6	28.6	359	5	Q9V7Q2
45	6	28.6	360	16	Q928L6
46	6	28.6	370	10	Q9LYK0
47	6	28.6	421	16	Q92WM5
48	6	28.6	425	10	Q9MA53
49	6	28.6	465	10	Q9FHT1
50	6	28.6	466	4	Q15846
51	6	28.6	472	3	Q9UU82
52	6	28.6	478	5	O9W0H5
53	6	28.6	484	5	O26475
54	6	28.6	496	2	O9AD76
55	6	28.6	499	3	O9C4B1
56	6	28.6	499	3	Q12551
57	6	28.6	506	3	Q96X02
58	6	28.6	506	3	Q96VA1
59	6	28.6	506	3	Q96VA0
60	6	28.6	507	4	Q96CL7
61	6	28.6	510	3	O74288
62	6	28.6	511	10	Q942V7
63	6	28.6	519	5	Q22017
64	6	28.6	534	5	O961B3
65	6	28.6	540	2	O68145
66	6	28.6	559	4	O9NVH1
67	6	28.6	561	2	O9EXF3
68	6	28.6	575	5	O45066
69	6	28.6	583	5	O61711
70	6	28.6	583	5	O27472
71	6	28.6	590	11	O88337
72	6	28.6	608	4	Q9NQX0
73	6	28.6	612	3	Q9US12
74	6	28.6	612	5	Q22779
75	6	28.6	626	11	O88810
76	6	28.6	635	4	Q969Y5
77	6	28.6	649	17	Q980R1
78	6	28.6	671	5	Q9TZK7
79	6	28.6	693	11	Q9WV66
80	6	28.6	726	16	Q9AAB8
81	6	28.6	737	2	O9EXQ5
82	6	28.6	787	5	Q9NSV0
83	6	28.6	804	3	O13768
84	6	28.6	872	4	Q9H6U7
85	6	28.6	872	4	Q96GN3
86	6	28.6	909	6	Q9TTT4
87	6	28.6	914	11	Q91W15
88	6	28.6	1054	10	Q9C698
89	6	28.6	1126	17	Q976A4

Q62146	mus musculus
Q940T7	mus musculus
O16314	caenorhabdi
Q95LH6	bos taurus
Q83184	mink cell f
Q9PLC4	chlamydia m
Q9EPW1	rattus norv
O15772	trypanosoma
O15777	trypanosoma
O84471	chlamydia t
Q18894	caenorhabdi
Q9RE46	clostridium
Q9K8F9	bacillus ha
Q9KDP6	bacillus ha
Q97F64	clostridium
Q9PB74	xyella fas
O01509	caenorhabdi
Q64106	rattus sp.
Q93UN6	vibrio vuln
Q91259	pseudomonas
P94910	methanosarc
Q9UM08	homo sapien
O83509	treponema p
O87519	escherichia
P78846	schizosacch
O9GRL0	leishmania
O96IH2	homo sapien
Q9V7Q2	drosophila
Q928L6	chlamydia p
Q9LYK0	arabidopsis
Q92WM5	rhizobium m
Q9MA53	arabidopsis
Q9FHT1	arabidopsis
Q15846	homo sapien
Q9UU82	schizosacch
O9W0H5	drosophila
O26475	schistocerc
O9AD76	streptomyce
O9C4B1	aspergillus
Q12551	aspergillus
Q96X02	penicillium
Q96VA1	aspergillus
Q96VA0	aspergillus
Q96CL7	homo sapien
O74288	emericaella
Q942V7	oryza sativ
Q22017	caenorhabdi
Q961B3	drosophila
O68145	burkholderi
O9NVH1	homo sapien
Q9EXF3	escherichia
O45066	caenorhabdi
O61711	caenorhabdi
Q27472	caenorhabdi
O88337	rattus norv
Q9NQX0	homo sapien
Q9US12	schizosacch
Q22779	caenorhabdi
O88810	rattus norv
Q969Y5	homo sapien
Q980R1	sulfolobus
Q9TZK7	caenorhabdi
Q9WV66	mus musculus
Q9AAB8	caulobacter
Q9EXQ5	clostridium
Q9NSV0	caenorhabdi
O13768	schizosacch
Q9H6U7	homo sapien
Q96GN3	homo sapien
Q9TTT4	oryctolagus
Q91W15	mus musculus
Q9C698	arabidopsis
Q976A4	sulfolobus

90 Q91V56 12 Q91V56 6 28.6 1145 12 Q91V56
91 Q22725 10 Q22725 6 28.6 1184 10 Q22725
92 Q9QR22 11 Q9QR22 6 28.6 1263 11 Q9QR22
93 Q9QV00 4 Q9QV00 6 28.6 1291 4 Q9QV00
94 Q9QPR1 4 Q9QPR1 6 28.6 1359 4 Q9QPR1
95 Q73697 13 Q73697 6 28.6 1559 13 Q73697
96 Q91W25 11 Q91W25 6 28.6 1602 11 Q91W25
97 Q91W25 11 Q91W25 6 28.6 1643 11 Q91W25
98 Q91W25 11 Q91W25 6 28.6 1688 11 Q91W25
99 Q92733 4 Q92733 6 28.6 1769 4 Q92733
100 Q95226 4 Q95226 6 28.6 1912 4 Q95226
101 Q9UHB1 4 Q9UHB1 6 28.6 1977 4 Q9UHB1
102 Q92327 11 Q92327 6 28.6 1981 11 Q92327
103 Q9J157 3 Q9J157 6 28.6 1985 3 Q9J157
104 P87112 3 P87112 6 28.6 2100 3 P87112
105 Q902A7 13 Q902A7 6 28.6 2168 13 Q902A7
106 Q92305 11 Q92305 6 28.6 2169 11 Q92305
107 Q99QY4 16 Q99QY4 6 28.6 2271 16 Q99QY4
108 Q92675 4 Q92675 6 28.6 2322 4 Q92675
109 Q9J1C1 11 Q9J1C1 6 28.6 2346 11 Q9J1C1
110 Q9W7R4 13 Q9W7R4 6 28.6 2590 13 Q9W7R4
111 Q9W7R4 13 Q9W7R4 6 28.6 2715 13 Q9W7R4
112 Q92LM3 16 Q92LM3 6 28.6 3194 16 Q92LM3
113 Q62524 5 Q62524 6 28.6 3512 5 Q62524
114 P73139 16 P73139 6 28.6 3972 16 P73139
115 Q9TRV9 5 Q9TRV9 5 23.8 25 6 Q9TRV9
116 Q91FE0 40 Q91FE0 5 23.8 40 12 Q91FE0
117 P70476 42 P70476 5 23.8 42 12 P70476
118 Q9J522 47 Q9J522 5 23.8 47 16 Q9J522
119 Q9F214 49 Q9F214 5 23.8 49 2 Q9F214
120 Q9DBD2 50 Q9DBD2 5 23.8 50 13 Q9DBD2
121 Q9P0F0 52 Q9P0F0 5 23.8 52 16 Q9P0F0
122 Q95063 53 Q95063 5 23.8 53 10 Q95063
123 Q9P0G1 53 Q9P0G1 5 23.8 53 16 Q9P0G1
124 Q93882 54 Q93882 5 23.8 54 5 Q93882
125 Q98539 54 Q98539 5 23.8 54 6 Q98539
126 Q9TBJ4 55 Q9TBJ4 5 23.8 55 8 Q9TBJ4
127 Q57575 56 Q57575 5 23.8 56 13 Q57575
128 Q96PS3 57 Q96PS3 5 23.8 57 4 Q96PS3
129 Q93998 57 Q93998 5 23.8 57 10 Q93998
130 Q41933 60 Q41933 5 23.8 60 10 Q41933
131 Q93877 61 Q93877 5 23.8 61 5 Q93877
132 Q93878 65 Q93878 5 23.8 65 5 Q93878
133 Q93609 66 Q93609 5 23.8 66 5 Q93609
134 Q920149 67 Q920149 5 23.8 67 8 Q920149
135 Q9ER13 67 Q9ER13 5 23.8 67 11 Q9ER13
136 Q9JY08 67 Q9JY08 5 23.8 67 16 Q9JY08
137 Q48131 68 Q48131 5 23.8 68 8 Q48131
138 Q98H86 70 Q98H86 5 23.8 70 16 Q98H86
139 Q9D877 74 Q9D877 5 23.8 74 12 Q9D877
140 Q9WUN3 75 Q9WUN3 5 23.8 75 11 Q9WUN3
141 Q40946 76 Q40946 5 23.8 76 10 Q40946
142 Q40947 76 Q40947 5 23.8 76 10 Q40947
143 Q40952 76 Q40952 5 23.8 76 10 Q40952
144 Q9RRN2 76 Q9RRN2 5 23.8 76 16 Q9RRN2
145 Q9ZGZ1 77 Q9ZGZ1 5 23.8 77 2 Q9ZGZ1
146 Q93MD7 77 Q93MD7 5 23.8 77 2 Q93MD7
147 Q9YPP9 78 Q9YPP9 5 23.8 78 12 Q9YPP9
148 Q9Z2U6 79 Q9Z2U6 5 23.8 79 11 Q9Z2U6
149 Q9PDY5 79 Q9PDY5 5 23.8 79 16 Q9PDY5
150 Q94364 80 Q94364 5 23.8 80 10 Q94364
151 Q92483 82 Q92483 5 23.8 82 12 Q92483
152 Q41082 83 Q41082 5 23.8 83 5 Q41082
153 Q93875 84 Q93875 5 23.8 84 17 Q93875
154 Q97UP8 84 Q97UP8 5 23.8 84 17 Q97UP8
155 Q94364 85 Q94364 5 23.8 85 5 Q94364
156 Q9B8X2 86 Q9B8X2 5 23.8 86 8 Q9B8X2
157 Q93879 87 Q93879 5 23.8 87 5 Q93879
158 Q93880 88 Q93880 5 23.8 88 5 Q93880
159 Q97588 88 Q97588 5 23.8 88 6 Q97588
160 Q9L199 89 Q9L199 5 23.8 89 10 Q9L199
161 Q95S05 89 Q95S05 5 23.8 89 5 Q95S05
162 Q9H8Z4 89 Q9H8Z4 5 23.8 89 16 Q9H8Z4
90 Q977A4 90 Q977A4 5 23.8 90 17 Q977A4

163 Q9IV56 5 23.8 163 Q9IV56
164 Q2725 5 23.8 164 Q2725
165 Q9QY22 5 23.8 165 Q9QY22
166 Q9UIV0 5 23.8 166 Q9UIV0
167 Q9UPR1 5 23.8 167 Q9UPR1
168 Q73697 5 23.8 168 Q73697
169 Q91W25 5 23.8 169 Q91W25
170 Q91W25 5 23.8 170 Q91W25
171 Q91W25 5 23.8 171 Q91W25
172 Q92733 5 23.8 172 Q92733
173 Q95226 5 23.8 173 Q95226
174 Q9UHB1 5 23.8 174 Q9UHB1
175 Q92327 5 23.8 175 Q92327
176 Q9J157 5 23.8 176 Q9J157
177 P87112 5 23.8 177 P87112
178 Q902A7 5 23.8 178 Q902A7
179 Q92305 5 23.8 179 Q92305
180 Q99QY4 5 23.8 180 Q99QY4
181 Q92675 5 23.8 181 Q92675
182 Q9J1C1 5 23.8 182 Q9J1C1
183 Q9W7R4 5 23.8 183 Q9W7R4
184 Q9W7R4 5 23.8 184 Q9W7R4
185 Q92LM3 5 23.8 185 Q92LM3
186 Q62524 5 23.8 186 Q62524
187 P73139 5 23.8 187 P73139
188 Q9TRV9 5 23.8 188 Q9TRV9
189 Q91FE0 5 23.8 189 Q91FE0
190 P70476 5 23.8 190 P70476
191 Q9J522 5 23.8 191 Q9J522
192 Q9F214 5 23.8 192 Q9F214
193 Q9DBD2 5 23.8 193 Q9DBD2
194 Q9P0F0 5 23.8 194 Q9P0F0
195 Q95063 5 23.8 195 Q95063
196 Q9P0G1 5 23.8 196 Q9P0G1
197 Q93882 5 23.8 197 Q93882
198 Q98539 5 23.8 198 Q98539
199 Q9TBJ4 5 23.8 199 Q9TBJ4
200 Q57575 5 23.8 200 Q57575
201 Q96PS3 5 23.8 201 Q96PS3
202 Q93998 5 23.8 202 Q93998
203 Q41933 5 23.8 203 Q41933
204 Q93877 5 23.8 204 Q93877
205 Q93878 5 23.8 205 Q93878
206 Q93609 5 23.8 206 Q93609
207 Q920149 5 23.8 207 Q920149
208 Q9ER13 5 23.8 208 Q9ER13
209 Q9JY08 5 23.8 209 Q9JY08
210 Q48131 5 23.8 210 Q48131
211 Q98H86 5 23.8 211 Q98H86
212 Q9D877 5 23.8 212 Q9D877
213 Q9WUN3 5 23.8 213 Q9WUN3
214 Q40946 5 23.8 214 Q40946
215 Q40947 5 23.8 215 Q40947
216 Q40952 5 23.8 216 Q40952
217 Q9RRN2 5 23.8 217 Q9RRN2
218 Q9ZGZ1 5 23.8 218 Q9ZGZ1
219 Q93MD7 5 23.8 219 Q93MD7
220 Q9YPP9 5 23.8 220 Q9YPP9
221 Q9Z2U6 5 23.8 221 Q9Z2U6
222 Q9PDY5 5 23.8 222 Q9PDY5
223 Q94364 5 23.8 223 Q94364
224 Q92483 5 23.8 224 Q92483
225 Q41082 5 23.8 225 Q41082
226 Q93875 5 23.8 226 Q93875
227 Q97UP8 5 23.8 227 Q97UP8
228 Q94364 5 23.8 228 Q94364
229 Q9B8X2 5 23.8 229 Q9B8X2
230 Q93879 5 23.8 230 Q93879
231 Q93880 5 23.8 231 Q93880
232 Q97588 5 23.8 232 Q97588
233 Q9L199 5 23.8 233 Q9L199
234 Q95S05 5 23.8 234 Q95S05
235 Q9H8Z4 5 23.8 235 Q9H8Z4

Q92LQ2 rhizobium m
Q01695 penicillium
Q9WDE3 human immun
Q5642 yersinia pe
Q8676 archaeoglob
Q9051 fundulus he
Q97S59 streptococ
Q8467 bacterioph
Q9YWS7 human immun
Q9YF9 streptococ
Q9AFU5 shigella fl
Q92LQ2 rhizobium m
Q9YD9 styela clav
Q93X17 pharbitis n
Q9A7J8 caulobacter
Q9PB2 campylobact
Q0547 saccharomyc
Q15941 bombyx mori
Q9ARW1 cicet ariet
Q71731 human immun
Q9JSA8 chlamydia p
Q9MFC3 beta vulgar
Q39084 arabidopsis
Q9JWB2 neisseria m
Q92G18 rickettsia
Q73985 pyrococcus
Q02562 acrolepiops
Q9XST8 canis famli
Q9X7N0 streptomyce
Q91881 diadromus p
Q91877 diprion pin
Q61977 mus musculu
Q9G9Y7 mus musculu
Q89271 carnation m
Q91153 notophthalm
Q41211 striga asia
Q58083 pyrococcus
Q9YMG9 lymantria d
Q9XJ5 rhizobium m
Q9XDB6 corynebacte
Q67530 aquifex aeo
Q58311 pyrococcus
Q91789 taenia ovis
Q95YV8 leishmania
Q41658 influenza a
Q91TV4 neisseria m
Q9V8Q4 drosophila m
Q914N7 pseudomonas
Q9XV31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosophila
Q9XJ66 drosophila
Q9XJ65 drosophila
Q9XJ64 drosophila
Q9XJ63 drosophila
Q9XJ62 drosophila
Q9XJ61 drosophila
Q9XJ60 drosophila
Q9XJ59 drosophila
Q9XJ58 drosophila
Q9XJ57 drosophila
Q9XJ56 drosophila
Q9XJ55 drosophila
Q9XJ54 drosophila
Q9XJ53 drosophila
Q9XJ52 drosophila
Q9XJ51 drosophila
Q9XJ50 drosophila
Q9XJ49 drosophila
Q9XJ48 drosophila
Q9XJ47 drosophila
Q9XJ46 drosophila
Q9XJ45 drosophila
Q9XJ44 drosophila
Q9XJ43 drosophila
Q9XJ42 drosophila
Q9XJ41 drosophila
Q9XJ40 drosophila
Q9XJ39 drosophila
Q9XJ38 drosophila
Q9XJ37 drosophila
Q9XJ36 drosophila
Q9XJ35 drosophila
Q9XJ34 drosophila
Q9XJ33 drosophila
Q9XJ32 drosophila
Q9XJ31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosophila
Q9XJ66 drosophila
Q9XJ65 drosophila
Q9XJ64 drosophila
Q9XJ63 drosophila
Q9XJ62 drosophila
Q9XJ61 drosophila
Q9XJ60 drosophila
Q9XJ59 drosophila
Q9XJ58 drosophila
Q9XJ57 drosophila
Q9XJ56 drosophila
Q9XJ55 drosophila
Q9XJ54 drosophila
Q9XJ53 drosophila
Q9XJ52 drosophila
Q9XJ51 drosophila
Q9XJ50 drosophila
Q9XJ49 drosophila
Q9XJ48 drosophila
Q9XJ47 drosophila
Q9XJ46 drosophila
Q9XJ45 drosophila
Q9XJ44 drosophila
Q9XJ43 drosophila
Q9XJ42 drosophila
Q9XJ41 drosophila
Q9XJ40 drosophila
Q9XJ39 drosophila
Q9XJ38 drosophila
Q9XJ37 drosophila
Q9XJ36 drosophila
Q9XJ35 drosophila
Q9XJ34 drosophila
Q9XJ33 drosophila
Q9XJ32 drosophila
Q9XJ31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosophila
Q9XJ66 drosophila
Q9XJ65 drosophila
Q9XJ64 drosophila
Q9XJ63 drosophila
Q9XJ62 drosophila
Q9XJ61 drosophila
Q9XJ60 drosophila
Q9XJ59 drosophila
Q9XJ58 drosophila
Q9XJ57 drosophila
Q9XJ56 drosophila
Q9XJ55 drosophila
Q9XJ54 drosophila
Q9XJ53 drosophila
Q9XJ52 drosophila
Q9XJ51 drosophila
Q9XJ50 drosophila
Q9XJ49 drosophila
Q9XJ48 drosophila
Q9XJ47 drosophila
Q9XJ46 drosophila
Q9XJ45 drosophila
Q9XJ44 drosophila
Q9XJ43 drosophila
Q9XJ42 drosophila
Q9XJ41 drosophila
Q9XJ40 drosophila
Q9XJ39 drosophila
Q9XJ38 drosophila
Q9XJ37 drosophila
Q9XJ36 drosophila
Q9XJ35 drosophila
Q9XJ34 drosophila
Q9XJ33 drosophila
Q9XJ32 drosophila
Q9XJ31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosophila
Q9XJ66 drosophila
Q9XJ65 drosophila
Q9XJ64 drosophila
Q9XJ63 drosophila
Q9XJ62 drosophila
Q9XJ61 drosophila
Q9XJ60 drosophila
Q9XJ59 drosophila
Q9XJ58 drosophila
Q9XJ57 drosophila
Q9XJ56 drosophila
Q9XJ55 drosophila
Q9XJ54 drosophila
Q9XJ53 drosophila
Q9XJ52 drosophila
Q9XJ51 drosophila
Q9XJ50 drosophila
Q9XJ49 drosophila
Q9XJ48 drosophila
Q9XJ47 drosophila
Q9XJ46 drosophila
Q9XJ45 drosophila
Q9XJ44 drosophila
Q9XJ43 drosophila
Q9XJ42 drosophila
Q9XJ41 drosophila
Q9XJ40 drosophila
Q9XJ39 drosophila
Q9XJ38 drosophila
Q9XJ37 drosophila
Q9XJ36 drosophila
Q9XJ35 drosophila
Q9XJ34 drosophila
Q9XJ33 drosophila
Q9XJ32 drosophila
Q9XJ31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosophila
Q9XJ66 drosophila
Q9XJ65 drosophila
Q9XJ64 drosophila
Q9XJ63 drosophila
Q9XJ62 drosophila
Q9XJ61 drosophila
Q9XJ60 drosophila
Q9XJ59 drosophila
Q9XJ58 drosophila
Q9XJ57 drosophila
Q9XJ56 drosophila
Q9XJ55 drosophila
Q9XJ54 drosophila
Q9XJ53 drosophila
Q9XJ52 drosophila
Q9XJ51 drosophila
Q9XJ50 drosophila
Q9XJ49 drosophila
Q9XJ48 drosophila
Q9XJ47 drosophila
Q9XJ46 drosophila
Q9XJ45 drosophila
Q9XJ44 drosophila
Q9XJ43 drosophila
Q9XJ42 drosophila
Q9XJ41 drosophila
Q9XJ40 drosophila
Q9XJ39 drosophila
Q9XJ38 drosophila
Q9XJ37 drosophila
Q9XJ36 drosophila
Q9XJ35 drosophila
Q9XJ34 drosophila
Q9XJ33 drosophila
Q9XJ32 drosophila
Q9XJ31 drosophila
Q9XJ30 drosophila
Q9XJ29 drosophila
Q9XJ28 drosophila
Q9XJ27 drosophila
Q9XJ26 drosophila
Q9XJ25 drosophila
Q9XJ24 drosophila
Q9XJ23 drosophila
Q9XJ22 drosophila
Q9XJ21 drosophila
Q9XJ20 drosophila
Q9XJ19 drosophila
Q9XJ18 drosophila
Q9XJ17 drosophila
Q9XJ16 drosophila
Q9XJ15 drosophila
Q9XJ14 drosophila
Q9XJ13 drosophila
Q9XJ12 drosophila
Q9XJ11 drosophila
Q9XJ10 drosophila
Q9XJ09 drosophila
Q9XJ08 drosophila
Q9XJ07 drosophila
Q9XJ06 drosophila
Q9XJ05 drosophila
Q9XJ04 drosophila
Q9XJ03 drosophila
Q9XJ02 drosophila
Q9XJ01 drosophila
Q9XJ00 drosophila
Q9XJ99 drosophila
Q9XJ98 drosophila
Q9XJ97 drosophila
Q9XJ96 drosophila
Q9XJ95 drosophila
Q9XJ94 drosophila
Q9XJ93 drosophila
Q9XJ92 drosophila
Q9XJ91 drosophila
Q9XJ90 drosophila
Q9XJ89 drosophila
Q9XJ88 drosophila
Q9XJ87 drosophila
Q9XJ86 drosophila
Q9XJ85 drosophila
Q9XJ84 drosophila
Q9XJ83 drosophila
Q9XJ82 drosophila
Q9XJ81 drosophila
Q9XJ80 drosophila
Q9XJ79 drosophila
Q9XJ78 drosophila
Q9XJ77 drosophila
Q9XJ76 drosophila
Q9XJ75 drosophila
Q9XJ74 drosophila
Q9XJ73 drosophila
Q9XJ72 drosophila
Q9XJ71 drosophila
Q9XJ70 drosophila
Q9XJ69 drosophila
Q9XJ68 drosophila
Q9XJ67 drosoph

236	5	23.8	133	16	006717	O06717 bacillus su	309	159	16	Q98HM7	Q98hm7 rhizobium l
237	5	23.8	134	2	Q68689	O68689 yersinia pe	310	159	16	Q922F3	Q92zf3 rhizobium m
238	5	23.8	135	2	Q9VBM9	Q9vbm9 drosophila	311	160	5	Q9VLG0	Q9vlg0 drosophila
239	5	23.8	136	11	Q99NC5	Q99nc5 rattus norv	312	161	15	Q91WT8	Q91wt8 human immun
240	5	23.8	137	11	Q99NC5	Q99nc5 mus musculu	313	163	4	Q9NWX7	Q9nwx7 homo sapien
241	5	23.8	138	4	Q15203	Q15203 homo sapien	314	165	10	Q9SAW7	Q9saw7 triticum tu
242	5	23.8	139	2	Q9XC36	Q9xc36 pseudomonas	315	165	10	Q9SAW6	Q9saw6 triticum ae
243	5	23.8	140	2	Q9F7W5	Q9f7w5 mycobacteri	316	165	17	Q96Z25	Q96z25 sulfolobus
244	5	23.8	141	17	Q59457	Q59457 pyrococcus	317	166	5	Q9VVI9	Q9vvi9 drosophila
245	5	23.8	142	5	Q964H9	Q964h9 hydractinia	318	166	16	Q97I03	Q97i03 clostridium
246	5	23.8	143	5	Q9L1P0	Q9l1p0 streptomyce	319	167	2	Q9RPP8	Q9rpp8 chlamydia t
247	5	23.8	144	5	Q9NAY6	Q9nay6 papilio gla	320	167	4	Q96FG7	Q96fg7 homo sapien
248	5	23.8	145	5	Q9BLK2	Q9blk2 drosophila	321	167	10	Q9XFG2	Q9xf2 arabidopsis
249	5	23.8	146	5	Q9BLK0	Q9blk0 drosophila	322	167	10	Q9SYX7	Q9syx7 arabidopsis
250	5	23.8	147	5	Q9BLK4	Q9blk4 drosophila	323	167	16	Q98120	Q98120 chlamydia t
251	5	23.8	148	5	Q9BLK3	Q9blk3 drosophila	324	168	4	Q96FU6	Q96fu6 homo sapien
252	5	23.8	149	5	Q9BLK1	Q9blk1 drosophila	325	168	8	Q96002	Q96002 uncultured
253	5	23.8	150	5	Q9BLJ9	Q9blj9 drosophila	326	168	10	Q9LTD9	Q9ltd9 arabidopsis
254	5	23.8	151	5	Q9BLJ8	Q9blj8 drosophila	327	168	10	Q9LTD9	Q9ltd9 arabidopsis
255	5	23.8	152	5	Q9BLJ7	Q9blj7 drosophila	328	169	5	Q9NPN1	Q9nfn1 schistosoma
256	5	23.8	153	10	Q9FS32	Q9fs32 helianthus	329	169	15	Q9QRP1	Q9qrpl human immun
257	5	23.8	154	2	Q9RI18	Q9ri18 yersinia pe	330	170	2	Q91336	Q91336 pseudomonas
258	5	23.8	155	11	Q9D116	Q9d116 mus musculu	331	170	5	Q9NKK7	Q9nkw7 patinopecte
259	5	23.8	156	2	Q9LBR4	Q9lbr4 clostridium	332	170	10	Q9LJ86	Q9lj86 arabidopsis
260	5	23.8	157	10	Q41050	Q41050 pisum sativ	333	171	5	Q62520	Q62520 caenorhabdi
261	5	23.8	158	11	Q88674	Q88674 mus musculu	334	172	4	Q9NQ47	Q9nq47 homo sapien
262	5	23.8	159	8	Q9MJA1	Q9mj11 rana nigrom	335	172	4	Q96LY1	Q96ly1 homo sapien
263	5	23.8	160	8	Q9MJ40	Q9mj40 rana porosa	336	173	8	Q98821	Q98821 epipedobate
264	5	23.8	161	8	Q9MJ39	Q9mj39 rana porosa	337	173	16	Q92XY4	Q92xy4 thizobium m
265	5	23.8	162	8	Q9MJ38	Q9mj38 rana porosa	338	174	5	Q9N474	Q9n474 caenorhabdi
266	5	23.8	163	16	Q9MJ37	Q9mj37 rana porosa	339	174	10	Q942M2	Q942m2 oryza sativ
267	5	23.8	164	16	Q9P1F9	Q9p1f9 campylobact	340	174	16	Q9PFR1	Q9pfr1 xylella fas
268	5	23.8	165	16	Q9KDJ9	Q9kdj9 bacillus ha	341	175	11	Q9D7R4	Q9d7r4 mus musculu
269	5	23.8	166	16	Q9CJP9	Q9cjp9 pasteurilla	342	176	2	Q9WXC4	Q9wxc4 pseudomonas
270	5	23.8	167	16	Q9A2M3	Q9a2m3 caulobacter	343	176	11	Q9DC76	Q9dc76 mus musculu
271	5	23.8	168	16	Q52450	Q52450 pseudanabae	344	177	2	Q9ADU4	Q9adu4 salmonella
272	5	23.8	169	5	Q18549	Q18549 lytechinus	345	177	16	Q9KSI2	Q9ksl2 vibrio chol
273	5	23.8	170	5	Q18551	Q18551 lytechinus	346	178	4	Q9H3V1	Q9h3v1 homo sapien
274	5	23.8	171	4	Q96L01	Q96l01 homo sapien	347	178	12	Q97602	Q97602 cowpox viru
275	5	23.8	172	16	Q9A5N4	Q9a5n4 caulobacter	348	180	3	Q9UTV0	Q9utv0 schizosacch
276	5	23.8	173	16	Q9F3L9	Q9f3l9 streptomyce	349	180	10	Q9S7S7	Q9s7s7 arabidopsis
277	5	23.8	174	12	Q9D2A9	Q9d2a9 mus musculu	350	180	16	Q9ZAK0	Q9zak0 helicobacte
278	5	23.8	175	12	Q91Z89	Q91z89 indian pean	351	180	17	Q9HKQ7	Q9hkk0 thermoplasm
279	5	23.8	176	16	Q9PEA7	Q9pea7 xylella fas	352	182	6	Q77772	Q77772 cryotolagus
280	5	23.8	177	16	Q914T0	Q914t0 pseudomonas	353	182	15	Q85641	Q85641 moloney mur
281	5	23.8	178	17	Q9UZB0	Q9uzb0 pyrococcus	354	184	10	Q9MB00	Q9mb00 brassica ca
282	5	23.8	179	4	Q9UDN5	Q9udn5 homo sapien	355	184	16	Q9HYV2	Q9hyv2 pseudomonas
283	5	23.8	180	16	Q98DB5	Q98db5 rhizobium l	356	184	16	Q9A2V6	Q9a2v6 caulobacter
284	5	23.8	181	17	Q59057	Q59057 pyrococcus	357	186	5	Q9VD65	Q9vd65 drosophila
285	5	23.8	182	5	Q21151	Q21151 caenorhabdi	358	186	10	Q9C9L1	Q9c9l1 arabidopsis
286	5	23.8	183	11	Q99MF7	Q99mf7 rattus norv	359	186	15	Q99F43	Q99f43 human immun
287	5	23.8	184	2	Q34417	Q34417 helicobacte	360	187	15	Q85640	Q85640 murine leuk
288	5	23.8	185	2	Q9R300	Q9r300 helicobacte	361	187	15	Q83401	Q83401 moloney mur
289	5	23.8	186	2	Q87748	Q87748 helicobacte	362	187	17	Q9YER3	Q9yer3 aeropyrum p
290	5	23.8	187	2	Q32645	Q32645 helicobacte	363	188	8	Q9G8U3	Q9g8u3 rhodomonas
291	5	23.8	188	2	Q32646	Q32646 helicobacte	364	189	11	Q9CZV1	Q9czv1 mus musculu
292	5	23.8	189	2	Q32647	Q32647 helicobacte	365	190	1	Q31079	Q31079 methanosarc
293	5	23.8	190	2	Q32648	Q32648 helicobacte	366	190	4	Q9UH44	Q9uh44 homo sapien
294	5	23.8	191	2	Q32649	Q32649 helicobacte	367	190	5	Q9XXA3	Q9xxa3 caenorhabdi
295	5	23.8	192	2	Q32650	Q32650 helicobacte	368	190	16	Q9RUK5	Q9ruk5 deinococcus
296	5	23.8	193	2	Q32651	Q32651 helicobacte	369	192	15	Q9DV18	Q9dv18 human immun
297	5	23.8	194	2	Q32652	Q32652 helicobacte	370	193	8	Q95512	Q95512 microhylla s
298	5	23.8	195	2	Q32653	Q32653 helicobacte	371	193	17	Q27380	Q27380 methanother
299	5	23.8	196	2	Q32654	Q32654 helicobacte	372	193	17	Q975X4	Q975x4 sulfolobus
300	5	23.8	197	2	Q32655	Q32655 helicobacte	373	195	5	Q20188	Q20188 caenorhabdi
301	5	23.8	198	16	P73080	P73080 synechocyst	374	195	17	Q97XB7	Q97xb7 sulfolobus
302	5	23.8	199	16	Q921D3	Q921d3 rickettsia	375	196	2	Q9X620	Q9x620 salmonella
303	5	23.8	200	17	Q58591	Q58591 pyrococcus	376	196	17	Q57856	Q57856 pyrococcus
304	5	23.8	201	13	Q90W55	Q90w55 gallus gall	377	197	4	Q9PLS3	Q9pls3 homo sapien
305	5	23.8	202	4	Q96DE1	Q96del1 homo sapien	378	197	6	Q9BXC9	Q9bgx9 macaca fasc
306	5	23.8	203	10	Q94KV2	Q94kv2 manihot esc	379	197	16	Q97JH5	Q97jh5 clostridium
307	5	23.8	204	13	Q9DG95	Q9dg95 pagrus majo	380	199	5	Q9N4S8	Q9n4s8 caenorhabdi
308	5	23.8	205	16	Q9K6C2	Q9k6c2 bacillus ha	381	200	16	Q92V49	Q92v49 rhizobium m

382	5	23.8	200	17	026617	026617 methanother	455	5	23.8	234	17	09HSL9	Q9hsl9 halobacteri
383	5	23.8	201	10	Q9EF62	Q9if62 arabidopsis	456	5	23.8	235	6	Q9UM71	Q28937 sus scrofa
384	5	23.8	201	16	Q9R37	Q9zr37 rhizobium m	457	5	23.8	236	4	Q9UM71	Q9um71 homo sapien
385	5	23.8	201	10	Q9C415	Q9cl15 arabidopsis	458	5	23.8	236	4	Q9FA7	Q96f77 homo sapien
386	5	23.8	202	11	Q61533	Q61533 mus musculus	459	5	23.8	236	10	Q9LH79	Q9lh79 arabidopsis
387	5	23.8	202	16	Q9JU16	Q9ju16 neisseria m	460	5	23.8	236	16	Q9RVR0	Q9rvr0 deinococcus
388	5	23.8	202	16	Q9CMK0	Q9cmk0 pasteurella	461	5	23.8	236	16	Q9ABG9	Q9abg9 caulobacter
389	5	23.8	203	2	Q9WXH1	Q9wxh1 alcaligenes	462	5	23.8	237	2	Q93RV7	Q93rv7 streptomyces
390	5	23.8	203	10	Q91VA3	Q91va3 arabidopsis	463	5	23.8	237	4	Q96AB8	Q96ab8 homo sapien
391	5	23.8	204	2	Q9Z436	Q9z436 pseudomonas	464	5	23.8	238	10	Q94IH4	Q94ih4 nanochlori
392	5	23.8	204	10	Q96446	Q96446 glycine max	465	5	23.8	238	10	Q94IH3	Q94ih3 nanochlori
393	5	23.8	205	4	Q9H3V3	Q9h3v3 homo sapien	466	5	23.8	238	10	Q94IH2	Q94ih2 nanochlori
394	5	23.8	206	2	Q9Z3X8	Q9z3x8 pseudomonas	467	5	23.8	238	10	Q94IH1	Q94ih1 nanochlori
395	5	23.8	206	2	Q9EWV6	Q9ewv6 streptomyces	468	5	23.8	238	10	Q94IH0	Q94ih0 nanochlori
396	5	23.8	206	16	Q9JRA2	Q9jra2 neisseria m	469	5	23.8	238	10	Q94IC9	Q94ig9 nanochlori
397	5	23.8	207	16	Q9HXH4	Q9hxx4 pseudomonas	470	5	23.8	238	10	Q94IC8	Q94ig8 chlorella k
398	5	23.8	207	16	Q92D89	Q92d89 listeria in	471	5	23.8	239	2	Q93DV4	Q93dv4 unidentified
399	5	23.8	210	5	Q9VGA9	Q9vga9 drosophila	472	5	23.8	239	4	Q9GJQ5	Q96jq5 homo sapien
400	5	23.8	210	16	Q9KNI2	Q9kni2 vibrio chol	473	5	23.8	239	4	Q9GJQ5	Q96jq5 drosophila
401	5	23.8	210	16	Q9K6F6	Q9k6f6 bacillus ha	474	5	23.8	240	12	P89299	P89299 yam mosaic
402	5	23.8	211	5	Q18549	Q18549 caenorhabdi	475	5	23.8	240	16	Q9P818	Q9pb18 xyella fas
403	5	23.8	211	5	Q76804	Q76804 drosophila	476	5	23.8	241	16	Q9AGV6	Q9agv6 caulobacter
404	5	23.8	211	5	Q9VEK7	Q9vek7 drosophila	477	5	23.8	242	2	Q93F79	Q93f79 xanthomonas
405	5	23.8	212	16	Q9ANH4	Q9anh4 bradyrhizob	478	5	23.8	242	10	Q92U42	Q92u42 arabidopsis
406	5	23.8	212	16	Q9PJS8	Q9pis8 chlamydia m	479	5	23.8	242	10	Q94OT2	Q94ot2 arabidopsis
407	5	23.8	213	13	Q57503	Q57503 sceloporus	480	5	23.8	243	5	Q76178	Q76178 drosophila
408	5	23.8	214	5	Q95V17	Q95v17 frankliniell	481	5	23.8	243	10	Q94ED9	Q94ed9 oryza sativ
409	5	23.8	214	10	Q9SSM9	Q9ssm9 arabidopsis	482	5	23.8	244	5	Q9N2T8	Q9n2t8 caenorhabdi
410	5	23.8	214	10	Q9M4Y0	Q9m4y0 phalaenopsi	483	5	23.8	244	5	Q20804	Q20804 caenorhabdi
411	5	23.8	215	10	Q81322	Q81322 arabidopsis	484	5	23.8	244	5	Q963K5	Q963k5 cercomonas
412	5	23.8	216	2	Q24739	Q24739 streptomyces	485	5	23.8	244	5	Q963K4	Q963k4 cercomonas
413	5	23.8	216	12	Q98223	Q98223 molluscum c	486	5	23.8	244	5	Q963K3	Q963k3 cercomonas
414	5	23.8	216	16	Q9PBY5	Q9pbys xyella fas	487	5	23.8	244	5	Q963K0	Q963k0 cercomonas
415	5	23.8	217	15	Q9OQZ7	Q9oqz7 human immun	488	5	23.8	244	10	Q94FN8	Q94fn8 chlorarachn
416	5	23.8	217	17	Q59599	Q59599 pyrococcus	489	5	23.8	244	10	Q94FN7	Q94fn7 lotharella
417	5	23.8	218	5	Q93848	Q93848 caenorhabdi	490	5	23.8	244	10	Q94FN6	Q94fn6 lotharella
418	5	23.8	219	17	Q9V982	Q9v982 acropyrum p	491	5	23.8	244	10	Q94FN5	Q94fn5 lotharella
419	5	23.8	219	13	Q93402	Q93402 carassius a	492	5	23.8	245	5	Q9GS44	Q9gs44 drosophila
420	5	23.8	220	4	Q9BX18	Q9bx18 homo sapien	493	5	23.8	245	5	Q9GS37	Q9gs37 drosophila
421	5	23.8	220	11	Q9D8E5	Q9d8e5 mus musculus	494	5	23.8	245	10	Q9FUV2	Q9fuv2 gonionomas
422	5	23.8	222	10	Q23951	Q23951 gossypium h	495	5	23.8	246	2	Q93LL8	Q93ll8 nostoc punc
423	5	23.8	222	10	Q9XH32	Q9xh32 spinacia ol	496	5	23.8	247	5	Q9GS48	Q9gs48 drosophila
424	5	23.8	222	17	Q97WC8	Q97wc8 sulfolobus	497	5	23.8	247	10	Q9LRC6	Q9lrc6 nicotiana t
425	5	23.8	222	17	Q96ZL7	Q96z17 sulfolobus	498	5	23.8	247	10	Q94C79	Q94c79 arabidopsis
426	5	23.8	224	6	Q9M2T3	Q9mzt3 megaderma l	499	5	23.8	247	15	Q40167	Q40167 human immun
427	5	23.8	224	10	Q9LTS2	Q9lts2 arabidopsis	500	5	23.8	248	2	Q936H6	Q936h6 staphylococ
428	5	23.8	224	12	Q9G916	Q9g916 human herpe	501	5	23.8	248	4	Q9H2W1	Q9h2w1 homo sapien
429	5	23.8	225	4	Q9HC76	Q9hc76 homo sapien	502	5	23.8	248	5	Q9GS40	Q9gs40 drosophila
430	5	23.8	225	4	Q96966	Q96p96 homo sapien	503	5	23.8	248	5	Q23373	Q23373 caenorhabdi
431	5	23.8	225	10	Q943A4	Q943a4 oryza sativ	504	5	23.8	248	16	Q9C1F1	Q9c1f1 lactococcus
432	5	23.8	225	11	Q99N05	Q99n05 mus musculus	505	5	23.8	249	4	Q9H2N3	Q9h2n3 homo sapien
433	5	23.8	226	2	Q9X869	Q9x869 streptomyces	506	5	23.8	249	5	Q9GS36	Q9gs36 drosophila
434	5	23.8	226	5	Q9VLK7	Q9vlk7 drosophila	507	5	23.8	250	3	Q00017	Q00017 aspergillus
435	5	23.8	226	11	Q9ES61	Q9es61 mus musculus	508	5	23.8	250	4	Q9BY19	Q9by19 homo sapien
436	5	23.8	226	11	Q9D3F6	Q9d3f6 mus musculus	509	5	23.8	250	5	Q9VU06	Q9vu06 drosophila
437	5	23.8	226	16	Q74140	Q74140 synecocyst	510	5	23.8	250	5	Q9NEV7	Q9nev7 caenorhabdi
438	5	23.8	227	8	Q98679	Q98679 pisum sativ	511	5	23.8	250	12	Q9YMH6	Q9ymh6 lymantria d
439	5	23.8	228	2	Q52867	Q52867 shewanella	512	5	23.8	251	16	Q916F4	Q916f4 pseudomonas
440	5	23.8	228	16	Q9CJ23	Q9cj23 lactococcus	513	5	23.8	252	2	Q9ZGG8	Q9zgg8 heliobacill
441	5	23.8	229	11	Q9D9F6	Q9d9f6 mus musculus	514	5	23.8	252	4	Q9NVT7	Q9nvt7 homo sapien
442	5	23.8	229	11	Q9EPQ0	Q9eqp0 mesocricetu	515	5	23.8	252	5	Q9GS43	Q9gs43 drosophila
443	5	23.8	229	13	Q90235	Q90235 ambystoma m	516	5	23.8	252	5	Q9GPH5	Q9gph5 callosobruc
444	5	23.8	230	3	Q24973	Q24973 giardia lam	517	5	23.8	252	5	Q95ZU0	Q95zu0 caenorhabdi
445	5	23.8	230	8	Q9TA24	Q9ta24 toxodonta a	518	5	23.8	253	5	Q95069	Q95069 drosophila
446	5	23.8	230	16	Q9CMX6	Q9cmx6 pasteurella	519	5	23.8	253	5	Q95074	Q95074 drosophila
447	5	23.8	231	3	Q9US15	Q9us15 schizosacch	520	5	23.8	253	5	Q95094	Q95094 drosophila
448	5	23.8	232	2	Q86620	Q86620 streptomyces	521	5	23.8	253	5	Q95070	Q95070 drosophila
449	5	23.8	232	6	Q95210	Q95210 oryctolagus	522	5	23.8	253	5	Q95072	Q95072 drosophila
450	5	23.8	233	10	Q9SMQ0	Q9sqw0 pleurochrys	523	5	23.8	253	5	Q95082	Q95082 drosophila
451	5	23.8	233	16	Q97I83	Q97i83 clostridium	524	5	23.8	253	5	Q95084	Q95084 drosophila
452	5	23.8	234	2	Q93T84	Q93t84 brucella ab	525	5	23.8	253	5	Q95089	Q95089 drosophila
453	5	23.8	234	11	Q9DCX4	Q9dcx4 mus musculus	526	5	23.8	253	5	Q95090	Q95090 drosophila
454	5	23.8	234	17	Q9V2G0	Q9v2g0 pyrococcus	527	5	23.8	253	5	Q95091	Q95091 drosophila

528	5	23.8	253	5	Q95096	Q95096 drosophila	601	5	23.8	269	3	Q9C2V5	Q9C2V5 dichotomocl
529	5	23.8	253	5	Q95093	Q95093 drosophila	602	5	23.8	269	3	Q9C2V4	Q9C2V4 dicranophor
530	5	23.8	253	5	Q9G339	Q9G339 drosophila	603	5	23.8	269	3	Q9C2V6	Q9C2V6 dissophora
531	5	23.8	253	16	Q9A220	Q9A220 caulobacter	604	5	23.8	269	3	Q9C2V2	Q9C2V2 echinospora
532	5	23.8	254	5	Q9G345	Q9G345 drosophila	605	5	23.8	269	3	Q9C2V3	Q9C2V3 ellisomyces
533	5	23.8	254	5	Q9G345	Q9G345 drosophila	606	5	23.8	269	3	Q9C2V1	Q9C2V1 fennellomyces
534	5	23.8	254	5	Q95RM5	Q95RM5 drosophila	607	5	23.8	269	3	Q9C2V7	Q9C2V7 gibbertella
535	5	23.8	254	10	Q9FG62	Q9FG62 arabidopsis	608	5	23.8	269	3	Q9C2V0	Q9C2V0 gongronella
536	5	23.8	255	2	Q9AL79	Q9AL79 bruceella su	609	5	23.8	269	3	Q9C2U4	Q9C2U4 halteromyces
537	5	23.8	255	5	P91946	P91946 drosophila	610	5	23.8	269	3	Q9C2U5	Q9C2U5 helicostylu
538	5	23.8	255	5	Q95075	Q95075 drosophila	611	5	23.8	269	3	Q9C2U3	Q9C2U3 hesseltinell
539	5	23.8	255	5	Q9G346	Q9G346 drosophila	612	5	23.8	269	3	Q9C2U6	Q9C2U6 hyphomucor
540	5	23.8	255	5	Q9VB13	Q9VB13 drosophila	613	5	23.8	269	3	Q9C2U2	Q9C2U2 kirkomyces
541	5	23.8	257	3	Q9PA43	Q9PA43 aspergillus	614	5	23.8	269	3	Q9C2S9	Q9C2S9 micromyces
542	5	23.8	257	16	Q97N93	Q97N93 streptococc	615	5	23.8	269	3	Q9C2T7	Q9C2T7 mortierella
543	5	23.8	257	16	Q92U50	Q92U50 rhizobium m	616	5	23.8	269	3	Q9C2T2	Q9C2T2 mortierella
544	5	23.8	258	2	Q99P23	Q99P23 shigella fl	617	5	23.8	269	3	Q9C2T0	Q9C2T0 mortierella
545	5	23.8	258	5	Q9VIT7	Q9VIT7 drosophila	618	5	23.8	269	3	Q9C2S5	Q9C2S5 mortierella
546	5	23.8	258	5	Q9G847	Q9G847 drosophila	619	5	23.8	269	3	Q9C2S4	Q9C2S4 mortierella
547	5	23.8	258	5	Q45005	Q45005 caenorhabdi	620	5	23.8	269	3	Q9C2T8	Q9C2T8 mucor amphi
548	5	23.8	258	11	Q9D7K5	Q9D7K5 mus musculu	621	5	23.8	269	3	Q9C2T6	Q9C2T6 mucor circi
549	5	23.8	259	5	Q9G541	Q9G541 drosophila	622	5	23.8	269	3	Q9C2T5	Q9C2T5 mucor hiema
550	5	23.8	259	5	Q9G538	Q9G538 drosophila	623	5	23.8	269	3	Q9C2T4	Q9C2T4 mucor indic
551	5	23.8	259	5	Q9G534	Q9G534 drosophila	624	5	23.8	269	3	Q9C2T1	Q9C2T1 mucor mucod
552	5	23.8	259	5	Q9G531	Q9G531 drosophila	625	5	23.8	269	3	Q9C2S8	Q9C2S8 rhizomucor
553	5	23.8	259	5	Q9G530	Q9G530 drosophila	626	5	23.8	269	3	Q9C2S7	Q9C2S7 mucor ramos
554	5	23.8	259	16	Q9XSD6	Q9XSD6 ceratitidis c	627	5	23.8	269	3	Q9C2S6	Q9C2S6 mucor recur
555	5	23.8	262	2	Q9JRK3	Q9JRK3 bacillus ha	628	5	23.8	269	3	Q9C2S5	Q9C2S5 mucor recur
556	5	23.8	262	2	Q9VHK2	Q9VHK2 lactococcus	629	5	23.8	269	3	Q9C2T9	Q9C2T9 mycotypha a
557	5	23.8	262	16	Q67451	Q67451 aquifex aeo	630	5	23.8	269	3	Q9C2T3	Q9C2T3 mycotypha m
558	5	23.8	263	5	Q9G542	Q9G542 drosophila	631	5	23.8	269	3	Q9C2T0	Q9C2T0 parasitella
559	5	23.8	263	5	Q9G532	Q9G532 drosophila	632	5	23.8	269	3	Q9C205	Q9C205 phascolumyc
560	5	23.8	263	5	Q9G530	Q9G530 drosophila	633	5	23.8	269	3	Q9C204	Q9C204 phycomyces
561	5	23.8	263	6	Q9XSD6	Q9XSD6 oryctolagus	634	5	23.8	269	3	Q9C207	Q9C207 pilaira ano
562	5	23.8	263	11	Q9D0M9	Q9D0M9 mus musculu	635	5	23.8	269	3	Q9C210	Q9C210 pilobolus u
563	5	23.8	263	12	Q89061	Q89061 variola vir	636	5	23.8	269	3	Q9C203	Q9C203 pirella cir
564	5	23.8	263	13	Q90P54	Q90P54 gallus gall	637	5	23.8	269	3	Q9C202	Q9C202 poitrassia c
565	5	23.8	264	2	Q9X647	Q9X647 aeromonas h	638	5	23.8	269	3	Q9C201	Q9C201 protomyccol
566	5	23.8	264	5	Q9G0W7	Q9G0W7 caenorhabdi	639	5	23.8	269	3	Q9C1Y2	Q9C1Y2 radiomyces
567	5	23.8	264	6	Q9TT19	Q9TT19 macaca fusc	640	5	23.8	269	3	Q9C1Y8	Q9C1Y8 rhizomucor
568	5	23.8	264	12	Q91S72	Q91S72 planaria as	641	5	23.8	269	3	Q9C1Y3	Q9C1Y3 rhizomucor
569	5	23.8	264	12	Q91S70	Q91S70 planaria as	642	5	23.8	269	3	Q9C1Y0	Q9C1Y0 rhizomucor
570	5	23.8	264	16	Q98MV9	Q98MV9 rhizobium l	643	5	23.8	269	3	Q9C1Y9	Q9C1Y9 rhizopus az
571	5	23.8	265	16	P74612	P74612 synecocyst	644	5	23.8	269	3	Q9C1Y7	Q9C1Y7 rhizopus mi
572	5	23.8	265	17	Q9YCO2	Q9YCO2 aeropyrum p	645	5	23.8	269	3	Q9C1Y6	Q9C1Y6 rhizopus mi
573	5	23.8	265	17	Q97ZJ5	Q97ZJ5 sulfobolus	646	5	23.8	269	3	Q9C1Y5	Q9C1Y5 rhizopus ol
574	5	23.8	266	2	Q93T29	Q93T29 burkholderi	647	5	23.8	269	3	Q9C1Y4	Q9C1Y4 rhizopus or
575	5	23.8	266	5	Q9VME6	Q9VME6 drosophila	648	5	23.8	269	3	Q9C1V8	Q9C1V8 saksanaea v
576	5	23.8	266	17	Q29396	Q29396 archaeoglob	649	5	23.8	269	3	Q9C1V9	Q9C1V9 spinellus f
577	5	23.8	267	4	Q9NXX0	Q9NXX0 homo sapien	650	5	23.8	269	3	Q9C1V7	Q9C1V7 syncephalas
578	5	23.8	268	11	Q9EQ20	Q9EQ20 mus musculu	651	5	23.8	269	3	Q9C1W0	Q9C1W0 syncephalas
579	5	23.8	268	11	Q99N10	Q99N10 mus musculu	652	5	23.8	269	3	Q9C1W0	Q9C1W0 synecocyst
580	5	23.8	269	3	Q9C494	Q9C494 absidia bla	653	5	23.8	269	3	Q9C1X8	Q9C1X8 syzygies m
581	5	23.8	269	3	Q9C493	Q9C493 absidia coe	654	5	23.8	269	3	Q9C1X5	Q9C1X5 thamnidiu
582	5	23.8	269	3	Q9C492	Q9C492 absidia cor	655	5	23.8	269	3	Q9C1V2	Q9C1V2 thamnostylu
583	5	23.8	269	3	Q9C2V1	Q9C2V1 absidia gla	656	5	23.8	269	3	Q9C1V3	Q9C1V3 thermomucor
584	5	23.8	269	3	Q9C2V8	Q9C2V8 absidia rep	657	5	23.8	269	3	Q9C1V0	Q9C1V0 umbelopsis
585	5	23.8	269	3	Q9C491	Q9C491 actinomucor	658	5	23.8	269	3	Q9C1U9	Q9C1U9 umbelopsis
586	5	23.8	269	3	Q9C2V7	Q9C2V7 amylomyces	659	5	23.8	269	3	Q9C1V1	Q9C1V1 utahomyces
587	5	23.8	269	3	Q9C490	Q9C490 apophysomyc	660	5	23.8	269	3	Q9C1U2	Q9C1U2 zychaea mex
588	5	23.8	269	3	Q9C2V3	Q9C2V3 backusella	661	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
589	5	23.8	269	3	Q9C2V2	Q9C2V2 backusella	662	5	23.8	269	3	Q9C1U2	Q9C1U2 zygorhynchu
590	5	23.8	269	3	Q9C2V0	Q9C2V0 benjaminie	663	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
591	5	23.8	269	3	Q9C2X9	Q9C2X9 blakeslea t	664	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
592	5	23.8	269	3	Q9C2X3	Q9C2X3 chaetocladi	665	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
593	5	23.8	269	3	Q9C2W0	Q9C2W0 chaetocladi	666	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
594	5	23.8	269	3	Q9C2V9	Q9C2V9 chlamydoabs	667	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
595	5	23.8	269	3	Q9C2X2	Q9C2X2 choanephora	668	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
596	5	23.8	269	3	Q9C2V7	Q9C2V7 circinella	669	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
597	5	23.8	269	3	Q9C153	Q9C153 circinomuco	670	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
598	5	23.8	269	3	Q9C2V8	Q9C2V8 cokeromyces	671	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
599	5	23.8	269	3	Q9C2X4	Q9C2X4 cunninghame	672	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu
600	5	23.8	269	3	Q9C2X1	Q9C2X1 cunninghame	673	5	23.8	269	3	Q9C1U3	Q9C1U3 zygorhynchu

674	273	10	023571	023571 arabidopsis	747	295	10	094BX8	094BX8 arabidopsis
675	273	10	09AV77	09av77 oryza sativ	748	295	11	035248	035248 cricetidae
676	273	13	098T06	098td6 paralichthy	749	295	16	091323	091323 pseudomonas
677	274	2	09R9K6	09r9k6 paracoccus	750	295	16	09CB40	09cb40 mycobacteri
678	274	5	020624	020624 chymomyza a	751	295	17	0974J0	0974j0 sulfolobus
679	274	5	020629	020629 chymomyza p	752	296	11	09CUR8	09cur8 mus musculus
680	274	5	020631	020631 drosophila	753	296	16	092SG7	092sg7 rhizobium m
681	274	5	020635	020635 zaprionus t	754	297	12	09QK06	09qk06 human enter
682	274	5	020630	020630 drosophila	755	297	16	09HYH2	09hyh2 pseudomonas
683	274	5	095080	095080 drosophila	756	297	16	092L10	092l10 rhizobium m
684	274	5	076807	076807 drosophila	757	297	16	092QV0	092qv0 rhizobium m
685	274	5	020633	020633 drosophila	758	298	16	09KD33	09kd33 bacillus ha
686	274	5	09NAH7	09nah7 caenorhabdi	759	298	16	094AV8	094av8 arabidopsis
687	275	2	09L332	09l332 serratia ma	760	300	2	09RJ57	09rj57 streptomyce
688	275	16	055193	055193 synecocyst	761	300	2	09L2H5	09l2h5 streptomyce
689	276	2	09RE15	09re15 klebsiella	762	300	4	096H23	096h23 homo sapien
690	276	5	09W5C7	09w5c7 drosophila	763	300	12	09ILC5	09ilc5 white spot
691	276	8	0951B9	0951b9 tetrahymena	764	300	12	0919R6	0919r6 white spot
692	276	10	09SDY0	09sdy0 lycopersico	765	300	16	092JW4	092jw4 rhizobium m
693	276	13	09W674	09w674 xenopus lae	766	301	16	033234	033234 mycobacteri
694	278	4	015509	015509 homo sapien	767	301	16	098174	098174 rhizobium l
695	278	11	0921K1	0921k1 mus musculu	768	301	16	092QV0	092qv0 rhizobium m
696	278	13	09PT89	09pt89 gallus gall	769	302	2	P72399	P72399 streptomyce
697	278	17	028227	028227 archaeglob	770	302	16	098DH2	098dh2 rhizobium l
698	279	5	0962P0	0962p0 schistosoma	771	303	2	030476	030476 bacillus su
699	279	10	09SDX9	09sdx9 lycopersico	772	303	2	09KZ08	09kz08 streptomyce
700	280	5	09XK99	09xk99 caenorhabdi	773	303	2	093EA6	093ea6 rhizobium l
701	280	10	09LIP8	09lip8 arabidopsis	774	303	5	09N6G9	09n6g9 leishmania
702	280	10	096405	096405 chlorella v	775	303	16	034543	034543 bacillus su
703	280	13	091152	091152 notophthalm	776	303	16	097G32	097g32 clostridium
704	280	16	092W65	092w65 rhizobium m	777	304	5	024110	024110 drosophila
705	281	13	091118	091118 morone saxa	778	305	2	087967	087967 streptomyce
706	281	13	090211	090211 sparus aura	779	305	11	09CWX0	09cw0 mus musculu
707	281	16	09CKS5	09cks5 pasteurella	780	305	16	09K9V6	09k9v6 bacillus ha
708	282	2	087877	087877 thauera aro	781	305	16	09JXM9	09jxm9 neisseria m
709	282	2	046980	046980 escherichia	782	305	16	09JW94	09jw94 neisseria m
710	282	10	09LJ24	09lj24 arabidopsis	783	306	2	097032	097032 streptococc
711	283	2	09RPW7	09rpw7 salmonella	784	306	2	054095	054095 saccharopol
712	283	3	096338	096338 saccharomyc	785	306	2	09XCL7	09xcl7 pseudomonas
713	283	5	09VBD0	09vbd0 drosophila	786	306	2	09AGZ6	09agz6 pseudomonas
714	284	5	0904H5	0904h5 drosophila	787	306	5	018513	018513 heliocidari
715	285	4	09H4T3	09h4t3 homo sapien	788	306	5	093253	093253 caenorhabdi
716	285	11	09ET48	09et48 marmota mon	789	306	10	09FTW9	09ftw9 oryza sativ
717	285	16	098HR2	098hr2 rhizobium l	790	306	16	09HZ90	09hz90 pseudomonas
718	286	2	09RKH4	09rk4 streptomyce	791	306	16	09HY52	09hy52 pseudomonas
719	286	5	09VER6	09ver6 drosophila	792	306	15	098C11	098c11 rhizobium l
720	286	5	045565	045565 caenorhabdi	793	306	15	0984Y9	0984y9 rhizobium l
721	286	16	0986C4	0986c4 rhizobium l	794	307	5	09GPK2	09gpk2 heterodera
722	287	2	P94923	P94923 mycobacteri	795	307	5	093531	093531 caenorhabdi
723	287	2	093A27	093a27 pseudomonas	796	307	16	09KRP7	09krp7 vibrio chol
724	287	10	09ZV50	09zv50 arabidopsis	797	307	16	09HW07	09hw07 pseudomonas
725	287	16	053755	053755 mycobacteri	798	309	16	09RS04	09rs04 deinococcus
726	287	16	09CJW6	09cjw6 pasteurella	799	311	5	09VDY7	09vdy7 drosophila
727	288	2	007487	007487 versinia en	800	311	10	09LNV7	09lnv7 arabidopsis
728	288	8	09B8Y5	09b8y5 fasciola he	801	311	16	0992G2	0992g2 streptococc
729	288	16	0992G1	0992g1 streptococc	802	312	5	062365	062365 caenorhabdi
730	288	17	058132	058132 pyrococcus	803	312	16	09K9D3	09k9d3 bacillus ha
731	291	5	09VG03	09vg03 drosophila	804	314	5	018860	018860 caenorhabdi
732	291	5	016260	016260 caenorhabdi	805	314	5	09U3R4	09u3r4 caenorhabdi
733	292	2	09ZBW3	09zbw3 streptomyce	806	314	5	09RI50	09ri50 caenorhabdi
734	292	5	09GNF2	09gnf2 dictyosteli	807	315	16	09KE32	09ke32 bacillus ha
735	292	5	09BML9	09bml9 helicospori	808	315	16	092VB3	092vb3 rhizobium m
736	292	16	099YQ9	099yq9 streptococc	809	316	4	09Y4R6	09y4r6 homo sapien
737	294	2	031094	031094 rhizobium l	810	316	16	067559	067559 aquifex aeo
738	294	2	09RMF3	09rmf3 zymomonas m	811	316	16	09PQI4	09pq14 ureaplasma
739	294	2	005699	005699 rhizobium l	812	316	16	09JVR1	09jvr1 neisseria m
740	294	11	09C9V8	09c9v8 mus musculu	813	316	16	098152	098152 rhizobium l
741	294	16	09RX39	09rx39 deinococcus	814	317	4	09NZR7	09nzt7 homo sapien
742	294	16	09JUN8	09jun8 neisseria m	815	317	5	09VBH7	09vbh7 drosophila
743	294	16	098K24	098k24 rhizobium l	816	317	5	046546	046546 mustela put
744	294	16	097523	097523 streptococc	817	317	16	P72026	P72026 mycobacteri
745	295	2	030530	030530 pseudomonas	818	317	16	092XJ0	092xj0 rhizobium m
746	295	2	050280	050280 pseudomonas	819	319	5	045006	045006 caenorhabdi


```

966 5 23.8 343 2 Q9X3Q9 Q9x3q9 pseudomonas
967 5 23.8 343 11 Q9J114 Q9j114 mus musculus
968 5 23.8 344 5 Q9XW59 Q9xw59 caenorhabdi
969 5 23.8 345 10 Q9SWP8 Q9swp8 pleurochrys
970 5 23.8 345 16 Q915H8 Q915h8 pseudomonas
971 5 23.8 346 5 Q9UAD9 Q9uad9 caenorhabdi
972 5 23.8 346 10 Q94FF0 Q94fp0 chlorarachn
973 5 23.8 346 10 Q94FN9 Q94fn9 chlorarachn
974 5 23.8 347 5 Q91700 Q91700 branchiost
975 5 23.8 347 5 Q9XW32 Q9xw32 caenorhabdi
976 5 23.8 347 10 Q9FUV3 Q9fuv3 cryptomonas
977 5 23.8 347 10 Q9C6X9 Q9c6x9 arabidopsis
978 5 23.8 348 5 Q9TGF4 Q9tgf4 caenorhabdi
979 5 23.8 348 5 Q91304 Q91304 caenorhabdi
980 5 23.8 348 5 Q92786 Q92786 caenorhabdi
981 5 23.8 349 11 Q91275 Q91275 mus musculu
982 5 23.8 350 2 Q9RKQ6 Q9rkq6 streptomyce
983 5 23.8 350 5 Q92757 Q92757 drosophila
984 5 23.8 350 5 Q9TVW7 Q9tvw7 drosophila
985 5 23.8 350 5 Q9UAY5 Q9uay5 caenorhabdi
986 5 23.8 350 10 Q9XFB2 Q9xfb2 lycopersico
987 5 23.8 350 16 Q97564 Q97564 bacillus su
988 5 23.8 351 3 Q94542 Q94542 schizosacch
989 5 23.8 351 10 Q9LFD7 Q9lfd7 arabidopsis
990 5 23.8 351 16 Q93837 Q93837 mycobacteri
991 5 23.8 351 16 Q97HY0 Q97hy0 clostridium
992 5 23.8 352 10 Q9XFB3 Q9xfb3 lycopersico
993 5 23.8 352 12 Q91BV2 Q91bv2 turkey herp
994 5 23.8 352 12 Q9S6P5 Q9s6p5 turkey herp
995 5 23.8 353 2 Q91AY8 Q91ay8 vibrio chol
996 5 23.8 353 2 Q9ZF09 Q9zf09 mycoplasma
997 5 23.8 353 5 Q9TYF9 Q9tyf9 drosophila
998 5 23.8 353 5 Q9TVF7 Q9tvf7 drosophila
999 5 23.8 353 5 Q9TVH7 Q9tvh7 drosophila
1000 5 23.8 353 5 Q92768 Q92768 drosophila

```

ALIGNMENTS

```

RESULT 1
ID Q99Q15 PRELIMINARY; PRT; 46 AA.
AC Q99Q15;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN PRECURSOR.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GB14, H7, LT11, NGS, AND UAI59;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in Biofilms.";
RT J. Bacteriol. 183:897-908(2001).
DR EMBL; AF277152; AAK01542.1; -.
DR EMBL; AF277153; AAK01543.1; -.
DR EMBL; AF277155; AAK01545.1; -.
DR EMBL; AF277157; AAK01546.1; -.
DR EMBL; AF277156; AAK01547.1; -.
DR InterPro; IPR004288; ComC.
DR Pfam; PF03047; ComC; 1.
FT CHAIN 26 46
SQ SEQUENCE 46 AA; 5211 MW; 38FA62B6F78FC3BF CRC64;

Query Match 100.0%; Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.1e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3
ID Q9APK6 PRELIMINARY; PRT; 43 AA.
AC Q9APK6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH1005;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in Biofilms.";
RT J. Bacteriol. 183:897-908(2001).
DR EMBL; AF277154; AAK01544.1; -.
FT CHAIN 26 43
SQ SEQUENCE 43 AA; 4927 MW; E6A78FC3BF6156C7 CRC64;

Query Match 71.4%; Score 15; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLFNRSFTQALGK 21
DB 26 SGSLSTFFRLFNRSFTQALGK 46

RESULT 2
ID Q9APK7 PRELIMINARY; PRT; 46 AA.
AC Q9APK7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BM71;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in Biofilms.";
RT J. Bacteriol. 183:897-908(2001).
DR EMBL; AF277151; AAK01541.1; -.
FT CHAIN 26 46
SQ SEQUENCE 46 AA; 5195 MW; 38E0B9DB5B8FC3BF CRC64;

Query Match 100.0%; Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.1e-15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGSLSTFFRLFNRSFTQALGK 21
DB 26 SGSLSTFFRLFNRSFTQALGK 46

RESULT 4
ID Q9APK6 PRELIMINARY; PRT; 43 AA.
AC Q9APK6;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE COMPETENCE STIMULATING PROTEIN.
GN COMC.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH1005;
RX MEDLINE=21142515; PubMed=11208787;
RA Li Y.H., Lau P.C.Y., Lee J.H., Ellen R.P., Cvitkovitch D.G.;
RT "Natural Genetic Transformation of Streptococcus mutans Growing in Biofilms.";
RT J. Bacteriol. 183:897-908(2001).
DR EMBL; AF277154; AAK01544.1; -.
FT CHAIN 26 43
SQ SEQUENCE 43 AA; 4927 MW; E6A78FC3BF6156C7 CRC64;

Query Match 71.4%; Score 15; DB 2; Length 43;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LSTFFRLFNRSFTQA 18
DB 29 LSTFFRLFNRSFTQA 43

```

```

RESULT 4
Q90315 ID Q90315 PRELIMINARY; PRT; 253 AA.
AC Q90315;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE STEM CELL FACTOR, MEMBRANE-BOUND FORM.
OS Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96283808; PubMed=8679698;
RA Pettit J.N., Kulik M.J.;
RT "Cloning and characterization of cDNAs encoding two forms of avian
RT stem cell factor.";
RL Biochim. Biophys. Acta 1307:149-151(1996).
DR EMBL; U43079; AAC59934.1; -.
DR InterPro; IPR003452; SCF.
DR Pfam; PF02404; SCF; 1.
SQ SEQUENCE 253 AA; 28993 MW; 20709B3854F0207A CRC64;

Query Match 33.3%; Score 7; DB 13; Length 253;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 13
DB 144 FFRLENR 150

RESULT 5
Q90314 ID Q90314 PRELIMINARY; PRT; 287 AA.
AC Q90314;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE STEM CELL FACTOR.
OS Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96283808; PubMed=8679698;
RA Pettit J.N., Kulik M.J.;
RT "Cloning and characterization of cDNAs encoding two forms of avian
RT stem cell factor.";
RL Biochim. Biophys. Acta 1307:149-151(1996).
DR EMBL; U43078; AAC59933.1; -.
DR InterPro; IPR003452; SCF.
DR Pfam; PF02404; SCF; 1.
SQ SEQUENCE 287 AA; 32455 MW; ABA81AEA422A702E CRC64;

Query Match 33.3%; Score 7; DB 13; Length 287;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 FFRLENR 13
DB 144 FFRLENR 150

RESULT 6
Q9VR32 ID Q9VR32 PRELIMINARY; PRT; 929 AA.

```

```

Q9VR32;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG15627 PROTEIN.
GN CG15627
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Flossler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J.D., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003575; AAF50976.1; -.
DR HSSP; P19491; IGR2.
DR FlyBase; FBgn0031634; CG15627.
DR InterPro; IPR001622; Channel_pore_K.
DR InterPro; IPR001320; Ion_glut_receptor.
DR InterPro; IPR001220; Lectin_legB.
DR InterPro; IPR001508; NMDA_receptor.
DR InterPro; IPR001311; SBP_glut_receptor.
DR Pfam; PF00060; lig_chan_1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PBPc; 1.
DR PROSITE; PS00307; LECTIN_LEGUME_BETA; UNKNOWN1.
SQ SEQUENCE 929 AA; 106132 MW; F4BF81C375A9CFED CRC64;

Query Match 33.3%; Score 7; DB 5; Length 929;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 SFTOALG 20
DB 144 SFTOALG 20

```


Db 103 SFTQALG 109

RESULT 7

Q1G53 PRELIMINARY; PRT; 51 AA.

AC Q1G53

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)

DE 046R.

OS Chilo iridescent virus (CIV) (Insect iridescent virus type 6).

OC Viruses; dsDNA viruses, no RNA stage; Iridoviridae; Iridovirus.

OX NCBI_TaxID=10488;

RN [1]

RP SEQUENCE FROM N.A.

RA Delliuss H., Darai G., Fluegel R.M.; evidence for circular

RT "DNA analysis of insect iridescent virus 6: evidence for circular

RL J. Virol. 49:609-614(1984).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=86174607; PubMed=3959991;

RA Lorbacher de Ruiz H., Gelderblom H., Hofmann W., Darai G.;

RT "Insect iridescent virus type 6 induced toxic degenerative hepatitis

RL in mice."

RN [3]

RP Med. Microbiol. Immunol. 175:43-53(1986).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=87321126; PubMed=2820141;

RA Schnitzler P., Soltan J.B., Fischer M., Relsner H., Scholz J.,

RA Delliuss H., Darai G.;

RT "Molecular cloning and physical mapping of the genome of insect

RL iridescent virus type 6: further evidence for circular permutation of

RL the viral genome."

RN [4]

RP Virology 160:66-74(1987).

RN [4]

RP SEQUENCE FROM N.A.

RX MEDLINE=89073752; PubMed=3201750;

RA Fischer M., Schnitzler P., Delliuss H., Darai G.;

RT "Identification and characterization of the repetitive DNA element in

RL the genome of insect iridescent virus type 6."

RN [5]

RP Virology 167:485-496(1988).

RN [5]

RP SEQUENCE FROM N.A.

RX MEDLINE=92196996; PubMed=1549908;

RA Handermann M., Schnitzler P., Rosen-Wolff A., Raab K., Sonntag K.C.,

RA Darai G.;

RT "Identification and mapping of origins of DNA replication within the

RL DNA sequences of the genome of insect iridescent virus type 6."

RN [6]

RP Virus Genes 6:19-32(1992).

RN [6]

RP SEQUENCE FROM N.A.

RX MEDLINE=93118242; PubMed=1475907;

RA Sonntag K.C., Darai G.;

RT "Characterization of the third origin of DNA replication of the genome

RL of insect iridescent virus type 6."

RN [7]

RP Virus Genes 6:333-342(1992).

RN [7]

RP SEQUENCE FROM N.A.

RX MEDLINE=93260401; PubMed=8492091;

RA Stonwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;

RT "Identification of the gene encoding the major capsid protein of

RL insect iridescent virus type 6 by polymerase chain reaction."

RN [8]

RP J. Gen. Virol. 74:873-879(1993).

RN [8]

RP SEQUENCE FROM N.A.

RX MEDLINE=94167241; PubMed=8121799;

RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,

RA Delliuss H., Darai G.;

RT "Identification of genes encoding zinc finger proteins, non-histone

RL chromosomal HMG protein homologue, and a putative GTP phosphohydrolase

RL in the genome of Chilo iridescent virus."

Nucleic Acids Res. 22:158-166(1994).

RN [9]

RP SEQUENCE FROM N.A.

RX MEDLINE=94353641; PubMed=8071636;

RA Sonntag K.C., Schnitzler P., Koonin E.V., Darai G.;

RT "Chilo iridescent virus encodes a putative helicase belonging to a

RL distinct family within the 'DEAD/H' superfamily: implications for the

RL evolution of large DNA viruses."

RN [10]

RP Virus Genes 8:151-158(1994).

RN [10]

RP SEQUENCE FROM N.A.

RX MEDLINE=95213160; PubMed=7698884;

RA Sonntag K.C., Schnitzler P., Janssen W., Darai G.;

RT "Identification of the primary structure and the coding capacity of

RL the genome of insect iridescent virus type 6 between the genome

RL coordinates 0.310 and 0.347 (7990 bp)."

RN [11]

RP Intervirology 37:287-297(1994).

RN [11]

RP SEQUENCE FROM N.A.

RX MEDLINE=94292906; PubMed=8021587;

RA Schnitzler P., Sonntag K.C., Muller M., Janssen W., Bugert J.J.,

RA Koonin E.V., Darai G.;

RT "Insect iridescent virus type 6 encodes a polypeptide related to the

RL largest subunit of eukaryotic RNA polymerase II."

RN [12]

RP J. Gen. Virol. 75:1557-1567(1994).

RN [12]

RP SEQUENCE FROM N.A.

RX MEDLINE=98141693; PubMed=9482589;

RA Bahr U., Tidona C.A., Darai G.;

RT "The DNA sequence of Chilo iridescent virus between the genome

RL coordinates 0.101 and 0.391; similarities in coding strategy between

RL insect and vertebrate iridoviruses."

RN [13]

RP Virus Genes 15:235-245(1997).

RN [13]

RP SEQUENCE FROM N.A.

RX MEDLINE=99125223; PubMed=9926400;

RA Muller K., Tidona C.A., Bahr U., Darai G.;

RT "Identification of a thymidylate synthase gene within the genome of

RL Chilo iridescent virus."

RN [14]

RP Virus Genes 17:243-258(1998).

RN [14]

RP SEQUENCE FROM N.A.

RX MEDLINE=99383793; PubMed=10456793;

RA Muller K., Tidona C.A., Darai G.;

RT "Identification of a gene cluster within the genome of Chilo

RL iridescent virus encoding enzymes involved in viral DNA replication

RL and processing."

RN [15]

RP Virus Genes 18:243-264(1999).

RN [15]

RP SEQUENCE FROM N.A.

RX MEDLINE=21342589; PubMed=11448171;

RA Jakob N.J., Muller K., Bahr U., Darai G.;

RT "Analysis of the First Complete DNA Sequence of an Invertebrate

RL Iridovirus: Coding Strategy of the Genome of Chilo Iridescent Virus."

RN [16]

RP Virology 286:182-196(2001).

RN [16]

RP SEQUENCE FROM N.A.

RX MEDLINE=93260401; PubMed=8492091;

RA Stonwasser R., Raab K., Schnitzler P., Janssen W., Darai G.;

RT "Identification of the gene encoding the major capsid protein of

RL insect iridescent virus type 6 by polymerase chain reaction."

RN [17]

RP J. Gen. Virol. 74:873-879(1993).

RN [17]

RP SEQUENCE FROM N.A.

RX MEDLINE=94167241; PubMed=8121799;

RA Schnitzler P., Hug M., Handermann M., Janssen W., Koonin E.V.,

RA Delliuss H., Darai G.;

RT "Identification of genes encoding zinc finger proteins, non-histone

RL chromosomal HMG protein homologue, and a putative GTP phosphohydrolase

RL in the genome of Chilo iridescent virus."

Query Match 28.6%; Score 6; DB 12; Length 51;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 SLSTFF 8
| | | | |
Db 35 SLSTFF 40

RESULT 8
041074

ID O41074 PRELIMINARY; PRT; 69 AA.
AC O41074;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE A592R PROTEIN.
GN A592R.
OS Paramacium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Phycodnavirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022962; PubMed=9356347;
RA Li Y., Lu Z., Sun L., Ropp S., Kutish G.F., Rock D.L., Van Etten J.L.;
RT "Analysis of 74 kb of DNA located at the right end of the 330-kb
RT chlorella virus PBCV-1 genome.";
RL Chlorella virus PBCV-1 genome.;
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20013326; PubMed=10544099;
RA Kaiser A., Vollmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Lisec A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homosperrmidine
RT synthase";
RL Virology 263:254-262(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT PBCV-1.";
RL Virology 276:27-36(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42580; AAC96934.1; .
SQ SEQUENCE 69 AA: 7861 MW; C3F24DE5A37F79B4 CRC64;
Query Match 28.6%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 SLSTFF 8
DB 27 SLSTFF 32
RESULT 9
Q91PP4 PRELIMINARY; PRT; 69 AA.
ID Q91PP4
AC Q91PP4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PRE-C PROTEIN (FRAGMENT).
GN PRE-C.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=97;
RA Ruifu Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=97;
RA Dai E.H., Song Y.J., Wang J., Liu H.J., Chen C.Y., Yang R.F.;
RT "Mutation specific PCR method for detecting nt1896 mutation in HBV
RT pre-C region.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB067454; BAB62297.1; .
FT NON_TER 1
SQ SEQUENCE 69 AA: 7702 MW; DAE075CC7F0C81DE CRC64;
Query Match 28.6%; Score 6; DB 12; Length 69;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 TFFRLF 11
DB 50 TFFRLF 55
RESULT 10
Q9RU75 PRELIMINARY; PRT; 122 AA.
ID Q9RU75
AC Q9RU75;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 13.8 KDA PROTEIN.
GN DRI518.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dadson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RL "Genome sequence of the radioresistant bacterium Deinococcus
RL radiodurans RL.";
RL Science 286:1571-1577(1999).
DR EMBL; AB001995; AAF11085.1; .
DR TIGR; DR1518; .
KW Hypothetical protein; Complete proteome
SQ SEQUENCE 122 AA: 13779 MW; 0F10C1F88C99E05B CRC64;
Query Match 28.6%; Score 6; DB 16; Length 122;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SGSLSLST 6
DB 66 SGSLSLST 71
RESULT 11

DT	01-JUN-2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	HYPOTHETICAL 15.2 KDA PROTEIN (FRAGMENT).
OS	Homo sapiens (Human).
OC	Eukaryota; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE-OVARY, ADENOCARCINOMA;
RA	Strausberg R.; to the EMBL/GenBank/DDBJ databases.
RC	Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RL	EMBL; BC005132; AAH05132.1; -.
DR	InterPro; IPR001871; bZIP.
DDR	SMART; SM00338; BRLZ; 1.
KW	Hypothetical protein.
FT	NON_TER 1
SEQ	SEQUENCE 140 AA; 15183 MW; 3A7658AA2FF4C7D CRC64;
Query Match	
Best Local Similarity 28.6%; Score 6; DB 4; Length 140;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 SGSLST 6
DDB	16 SGSLST 21
RESULT 13	
Q9ORM3	PRELIMINARY; PRT; 177 AA.
ID	Q9ORM3
AC	Q9ORM3
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE	ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN	ENV.
OS	Human immunodeficiency virus type 1.
OS	Viruses; Retroid viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11676;	[1]
RN	SEQUENCE FROM N.A.
RP	STRAIN=96G38_9;
EC	MEGLINE=20250533; PubMed=10791871;
EX	Bikandou B., Takehisa J., Mboudjeka I., Ido E., Kuwata T.,
RA	Miyazaki Y., Moriyama H., Harada Y., Taniguchi Y., Ichimura H.,
RA	Ikedo M., Ndolo P.J., Nzoukoudi M.Y., M'Vouenze R., M'Pandi M.,
RA	Perra H.J., M'Pele P., Hayami M.;
RA	AIDS Res. Hum. Retroviruses 16:613-619(2000)."
FT	Non_TER 1
SEQ	SEQUENCE 177 AA; 20456 MW; D7BF88A336CA5ED4 CRC64;
Query Match	
Best Local Similarity 28.6%; Score 6; DB 15; Length 177;	
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	9 RLFNRS 14
DDB	148 RLFNRS 153
RESULT 14	
Q98FXO	PRELIMINARY; PRT; 180 AA.
ID	Q98FXO
AC	Q98FXO;
DT	01-OCT-2001 (TrEMBLrel. 18, Created)

```

DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
GN MLR3584 PROTEIN.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003002; BAB50446.1; -.
KW Complete proteome.
SQ SEQUENCE 180 AA; 20202 MW; 054EDA7E73943DB6 CRC64;

Query Match      28.6%; Score 6; DB 16; Length 180;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RSFTQA 18
   |||||
Db 61 RSFTQA 66

RESULT 15
Q901F3 Q901F3 PRELIMINARY; PRT; 181 AA.
AC Q901F3;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ENVELOPE GLYCOPROTEIN (FRAGMENT).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=18006;
RA Holguin A., Alvarez A., Soriano V.;
RT "HIV-1 subtype J in Spain.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY050185; AAL13161.1; -.
FT NON_TER 1
FT NON_TER 181
SQ SEQUENCE 181 AA; 19870 MW; DBD76971DEE5FC51 CRC64;

Query Match      28.6%; Score 6; DB 15; Length 181;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LFNRSF 15
   |||||
Db 175 LFNRSF 180

Search completed: November 5, 2002, 11:06:58
Job time : 43.791 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 20:31:42 ; Search time 830.284 Seconds
(without alignments)
736.085 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 104

Sequence: 1 SGLSTFFRLNRSFTQALGK 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O/cn2_1/USPTO-Spool/US09833017/runat_05112002_105348_4834/app_query.fasta_1.398
-Db=GenEmbl -QWMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017 @CN_1.1.1616 @runat_05112002_105348_4834 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.om:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

29: em.vi:*
30: em.htg_hum:*
31: em.htg_inv:*
32: em.htg_mus:*
33: em.htg_other:*
34: em.htg_pln:*
35: em.htg_rod:*
36: em.htg_mam:*
37: em.htg_vrt:*
38: em.sy:*
39: em.htgo_hum:*
40: em.htgo_mus:*
41: em.htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total_score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	104	100.0	141	1	AF277151	AF277151 Streptoco
2	104	100.0	141	1	AF277152	AF277152 Streptoco
3	104	100.0	141	1	AF277153	AF277153 Streptoco
4	104	100.0	141	1	AF277155	AF277155 Streptoco
5	104	100.0	141	1	AF277156	AF277156 Streptoco
6	104	100.0	141	1	AF277157	AF277157 Streptoco
7	88	84.6	159	1	AF277154	AF277154 Streptoco
C 8	58	55.8	186403	9	AC022558	AC022558 Homo sapi
9	54	51.9	55713	9	AC073088	AC073088 Homo sapi
10	54	51.9	141605	2	AC013732	AC013732 Homo sapi
11	54	51.9	176203	2	AC094338	AC094338 Rattus no
C 12	54	51.9	231995	2	AC115352	AC115352 Rattus no
C 13	53	51.0	2613	14	WTVDI	J03020 Wound tumor
C 14	53	51.0	133226	2	AL512364	AL512364 Homo sapi
C 15	53	51.0	135820	9	AL513497	AL513497 Human DNA
C 16	53	51.0	149216	9	AL450338	AL450338 Human DNA
C 17	53	51.0	170203	2	AL714002	AL714002 Homo sapi
C 18	53	51.0	181112	2	AL360012	AL360012 Homo sapi
C 19	53	51.0	230668	2	AC121921	AC121921 Mus muscu
C 20	52	50.0	10475	6	I08101	I08101 Sequence 9
C 21	52	50.0	10475	6	I08101	I08101 Sequence 9
C 22	52	50.0	10476	6	AX334892	AX334892 Sequence
C 23	52	50.0	10476	6	AX335346	AX335346 Sequence
C 24	52	50.0	10476	9	HUMSPBAA	M24461 Human pulmo
C 25	52	50.0	11807	9	AF400074	AF400074 Homo sapi
C 26	52	50.0	54157	8	T7A14	AC005322 Arabidops
C 27	52	50.0	62001	2	AC015206	AC015206 Drosophill
C 28	52	50.0	99006	2	AC096254	AC096254 Rattus no
C 29	52	50.0	153094	2	AC096715	AC096715 Homo sapi
C 30	52	50.0	163174	2	AC103031	AC103031 Rattus no
C 31	52	50.0	164659	2	AC104143	AC104143 Drosophill
C 32	52	50.0	169189	2	AC125293	AC125293 Drosophill
C 33	52	50.0	185512	9	AC012454	AC012454 Homo sapi
C 34	52	50.0	222472	2	AC026091	AC026091 Homo sapi
C 35	52	50.0	229155	2	AC067301	AC067301 Homo sapi
C 36	52	50.0	310909	3	AE003427	AE003427 Drosophill
C 37	51.5	49.5	141659	2	AC078864	AC078864 Homo sapi
C 38	51.5	49.5	158285	9	AC021066	AC021066 Homo sapi
C 39	51.5	49.5	166046	2	AC080174	AC080174 Homo sapi
C 40	51	49.0	104117	9	AL591591	AL591591 Human DNA
C 41	51	49.0	142561	2	AC131137	AC131137 Rattus no
C 42	51	49.0	144046	2	AC068334	AC068334 Homo sapi
C 43	51	49.0	149532	2	AC123515	AC123515 Oryza sat
C 44	51	49.0	151131	2	AC021945	AC021945 Homo sapi
C 45	51	49.0	157581	2	AP005624	AP005624 Homo sapi

ALIGNMENTS

RESULT 1

AF277151 141 bp DNA linear BCT 07-FEB-2001
 LOCUS Streptococcus mutans strain BM71 competence stimulating protein precursor (comC) gene, complete cds.
 ACCESSION AF277151
 VERSION AF277151.1 GI:12698427
 KEYWORDS Streptococcus mutans.
 SOURCE Streptococcus mutans.
 ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
 TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
 JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)
 MEDLINE 21142515
 PUBMED 11208787
 REFERENCE 2 (bases 1 to 141)
 AUTHORS Lau, P.C.Y. and Cvitkovitch, D.G.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada
 FEATURES
 source
 1. .141
 /organism="Streptococcus mutans"
 /strain="BM71"
 /db_xref="taxon:1309"
 1. .141
 /gene="comC"
 1. .141
 /gene="comC"
 /note="comC"
 /codon_start=1
 /transl_table=11
 /product="competence stimulating protein precursor"
 /protein_id="AAK01541.1"
 /db_xref="GI:12698428"
 /translation="MKKTPSLKNDKFKEIKTDELEIIIGSGSLSTFFRLNRSFTQAL GK"
 76. .138
 /gene="comC"
 /product="competence stimulating protein"
 58 a 20 c 25 g 38 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.93e-09 Length: 141
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-833-017B-4 (1-21) x AF277151 (1-141)
 QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 76 ACCGGAAGCCTATCAACATTTTCGGCTGTTTACACAGAGTTTACACAGCTTTGGGA 135
 QY 21 Lys 21
 Db 136 AAA 138
 RESULT 2
 LOCUS AF277152 141 bp DNA linear BCT 07-FEB-2001
 DEFINITION Streptococcus mutans strain GB14 competence stimulating protein precursor (comC) gene, complete cds.
 ACCESSION AF277152
 VERSION AF277152.1 GI:12698429
 KEYWORDS Streptococcus mutans.
 SOURCE Streptococcus mutans.
 ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
 TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
 JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)
 MEDLINE 21142515
 PUBMED 11208787
 REFERENCE 2 (bases 1 to 141)
 AUTHORS Lau, P.C.Y. and Cvitkovitch, D.G.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada
 FEATURES
 source
 1. .141
 /organism="Streptococcus mutans"
 /strain="GB14"
 /db_xref="taxon:1309"
 1. .141
 /gene="comC"
 1. .141
 /gene="comC"
 /note="comC"
 /codon_start=1
 /transl_table=11
 /product="competence stimulating protein precursor"
 /protein_id="AAK01541.1"
 /db_xref="GI:12698428"
 /translation="MKKTPSLKNDKFKEIKTDELEIIIGSGSLSTFFRLNRSFTQAL GK"
 76. .138
 /gene="comC"
 /product="competence stimulating protein"
 58 a 20 c 25 g 38 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.93e-09 Length: 141
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-833-017B-4 (1-21) x AF277151 (1-141)
 QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 76 ACCGGAAGCCTATCAACATTTTCGGCTGTTTACACAGAGTTTACACAGCTTTGGGA 135
 QY 21 Lys 21
 Db 136 AAA 138
 RESULT 2
 LOCUS AF277152 141 bp DNA linear BCT 07-FEB-2001
 DEFINITION Streptococcus mutans strain GB14 competence stimulating protein precursor (comC) gene, complete cds.
 ACCESSION AF277152
 VERSION AF277152.1 GI:12698429
 KEYWORDS Streptococcus mutans.
 SOURCE Streptococcus mutans.

ORGANISM Streptococcus mutans
 Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
 TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
 JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)
 MEDLINE 21142515
 PUBMED 11208787
 REFERENCE 2 (bases 1 to 141)
 AUTHORS Lau, P.C.Y. and Cvitkovitch, D.G.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUN-2000) Microbiology, Dental Research Institute, 124 Edward Street, Toronto, ON M5G 1G6, Canada
 FEATURES
 source
 1. .141
 /organism="Streptococcus mutans"
 /strain="GB14"
 /db_xref="taxon:1309"
 1. .141
 /gene="comC"
 1. .141
 /gene="comC"
 /note="comC"
 /codon_start=1
 /transl_table=11
 /product="competence stimulating protein precursor"
 /protein_id="AAK01542.1"
 /db_xref="GI:12698430"
 /translation="MKKTI SLKNDKFKEIKTDELEIIIGSGSLSTFFRLNRSFTQAL GK"
 76. .138
 /gene="comC"
 /product="competence stimulating protein"
 58 a 19 c 25 g 39 t
 BASE COUNT
 ORIGIN
 Alignment Scores:
 Pred. No.: 7.93e-09 Length: 141
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 1 Gaps: 0
 US-09-833-017B-4 (1-21) x AF277152 (1-141)
 QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 76 ACCGGAAGCCTATCAACATTTTCGGCTGTTTACACAGAGTTTACACAGCTTTGGGA 135
 QY 21 Lys 21
 Db 136 AAA 138
 RESULT 3
 LOCUS AF277153 141 bp DNA linear BCT 07-FEB-2001
 DEFINITION Streptococcus mutans strain H7 competence stimulating protein precursor (comC) gene, complete cds.
 ACCESSION AF277153
 VERSION AF277153.1 GI:12698431
 KEYWORDS Streptococcus mutans.
 SOURCE Streptococcus mutans.
 ORGANISM Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 Streptococcus.
 REFERENCE 1 (bases 1 to 141)
 AUTHORS Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
 TITLE Natural genetic transformation of Streptococcus mutans growing in biofilms
 JOURNAL J. Bacteriol. 183 (3), 897-908 (2001)

```

MEDLINE      21142515
PUBMED      11208787
REFERENCE    2 (bases 1 to 141)
AUTHORS      Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE        Direct Submission
JOURNAL      Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
              124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES     Location/Qualifiers
source       1..141
              /organism="Streptococcus mutans"
              /strain="H7"
              /db_xref="taxon:1309"
gene         1..141
              /gene="comC"
CDS          1..141
              /gene="comC"
              /note="comC"
              /codon_start=1
              /transl_table=11
              /product="competence stimulating protein precursor"
              /protein_id="AAK01543.1"
              /db_xref="GI:12698432"
              /translation="MKKTLKNDKFKETDELEIIIGSGSLSTFFRLNRSFTQAL
              GK"
mat_peptide  76..138
              /gene="comC"
              /product="competence stimulating protein"

BASE COUNT  58 a 19 c 25 g 39 t
ORIGIN
Alignment Scores:
Pred. No.:      7,93e-09      Length:      141
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              1          Gaps:      0

US-09-833-017B-4 (1-21) x AF277153 (1-141)

QY  1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
    |||||
Db  76 AGCGGAGGCTATCAACATTTTCCGGCTGTTTACAGAAGTTTACACAAGCTTTGGGA 135

QY  21 Lys 21
    |||
Db  136 AAA 138

RESULT 4
AF277155
LOCUS      AF277155
DEFINITION Streptococcus mutans strain LT11 competence stimulating protein
            precursor (comC) gene, complete cds.
ACCESSION  AF277155
VERSION     AF277155.1 GI:12698435
KEYWORDS   J. Bacteriol. 183 (3), 897-908 (2001)
SOURCE     Streptococcus mutans.
           Streptococcus mutans
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
           Streptococcus.
REFERENCE  1 (bases 1 to 141)
AUTHORS    Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
TITLE      Natural genetic transformation of Streptococcus mutans growing in
           biofilms
JOURNAL    J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE    21142515
PUBMED     11208787
REFERENCE  2 (bases 1 to 141)
AUTHORS    Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE      Direct Submission
JOURNAL    Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
           124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES     Location/Qualifiers
source       1..141
              /organism="Streptococcus mutans"
              /strain="H7"
              /db_xref="taxon:1309"
gene         1..141
              /gene="comC"
CDS          1..141
              /gene="comC"
              /note="comC"
              /codon_start=1
              /transl_table=11
              /product="competence stimulating protein precursor"
              /protein_id="AAK01543.1"
              /db_xref="GI:12698432"
              /translation="MKKTLKNDKFKETDELEIIIGSGSLSTFFRLNRSFTQAL
              GK"
mat_peptide  76..138
              /gene="comC"
              /product="competence stimulating protein"

BASE COUNT  58 a 19 c 25 g 39 t
ORIGIN
Alignment Scores:
Pred. No.:      7,93e-09      Length:      141
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              1          Gaps:      0

US-09-833-017B-4 (1-21) x AF277153 (1-141)

QY  1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
    |||||
Db  76 AGCGGAGGCTATCAACATTTTCCGGCTGTTTACAGAAGTTTACACAAGCTTTGGGA 135

QY  21 Lys 21
    |||
Db  136 AAA 138

RESULT 5
AF277156
LOCUS      AF277156
DEFINITION Streptococcus mutans strain NG8 competence stimulating protein
            precursor (comC) gene, complete cds.
ACCESSION  AF277156
VERSION     AF277156.1 GI:12698437
KEYWORDS   J. Bacteriol. 183 (3), 897-908 (2001)
SOURCE     Streptococcus mutans.
           Streptococcus mutans
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
           Streptococcus.
REFERENCE  1 (bases 1 to 141)
AUTHORS    Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
TITLE      Natural genetic transformation of Streptococcus mutans growing in
           biofilms
JOURNAL    J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE    21142515
PUBMED     11208787
REFERENCE  2 (bases 1 to 141)
AUTHORS    Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE      Direct Submission
JOURNAL    Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
           124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES     Location/Qualifiers
source       1..141
              /organism="Streptococcus mutans"
              /strain="NG8"
              /db_xref="taxon:1309"
gene         1..141
              /gene="comC"
CDS          1..141
              /gene="comC"

```

```

source       1..141
              /organism="Streptococcus mutans"
              /strain="LT11"
              /db_xref="taxon:1309"
gene         1..141
              /gene="comC"
CDS          1..141
              /gene="comC"
              /note="comC"
              /codon_start=1
              /transl_table=11
              /product="competence stimulating protein precursor"
              /protein_id="AAK01545.1"
              /db_xref="GI:12698436"
              /translation="MKKTLKNDKFKETDELEIIIGSGSLSTFFRLNRSFTQAL
              GK"
mat_peptide  76..138
              /gene="comC"
              /product="competence stimulating protein"

BASE COUNT  58 a 19 c 25 g 39 t
ORIGIN
Alignment Scores:
Pred. No.:      7,93e-09      Length:      141
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    100.00%      Indels:      0
DB:              1          Gaps:      0

US-09-833-017B-4 (1-21) x AF277155 (1-141)

QY  1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
    |||||
Db  76 AGCGGAGGCTATCAACATTTTCCGGCTGTTTACAGAAGTTTACACAAGCTTTGGGA 135

QY  21 Lys 21
    |||
Db  136 AAA 138

RESULT 5
AF277156
LOCUS      AF277156
DEFINITION Streptococcus mutans strain NG8 competence stimulating protein
            precursor (comC) gene, complete cds.
ACCESSION  AF277156
VERSION     AF277156.1 GI:12698437
KEYWORDS   J. Bacteriol. 183 (3), 897-908 (2001)
SOURCE     Streptococcus mutans.
           Streptococcus mutans
ORGANISM   Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
           Streptococcus.
REFERENCE  1 (bases 1 to 141)
AUTHORS    Li, Y.H., Lau, P.C., Lee, J.H., Ellen, R.P. and Cvitkovitch, D.G.
TITLE      Natural genetic transformation of Streptococcus mutans growing in
           biofilms
JOURNAL    J. Bacteriol. 183 (3), 897-908 (2001)
MEDLINE    21142515
PUBMED     11208787
REFERENCE  2 (bases 1 to 141)
AUTHORS    Lau, P.C.Y. and Cvitkovitch, D.G.
TITLE      Direct Submission
JOURNAL    Submitted (13-JUN-2000) Microbiology, Dental Research Institute,
           124 Edward Street, Toronto, ON M5G 1G6, Canada
FEATURES     Location/Qualifiers
source       1..141
              /organism="Streptococcus mutans"
              /strain="NG8"
              /db_xref="taxon:1309"
gene         1..141
              /gene="comC"
CDS          1..141
              /gene="comC"

```


Percent Similarity:	77.78%	Conservative:	1
Best Local Similarity:	74.07%	Mismatches:	0
Query Match:	Indels:	Gaps:	6
DB:	1		1
US-09-833-017B-4 (1-21) x AF277154 (1-159)			
QY	1	SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSer	----- 14
Ddb	76	ACGGAGACCTGTCACACTTTTCGGCTGTGTTACAGAGATTTTACACAGCTTAGAAT	135
QY	15	PheThrGlnAlaLeuGlyLys 21	
Ddb	136	TTTACACAGCTTTGGGAAA 156	
RESULT 8			
AC022558/C			
LOCUS	AC022558	186403 bp	linear PRI 01-DEC-2001
DEFINITION	Homo sapiens chromosome 15, clone RP11-90B9, complete sequence.		
ACCESSION	AC022558		
VERSION	AC022558.9	GI:17223344	
KEYWORDS	HTG.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferrellaro,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczy,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,J., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Oliviar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,B., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.			
Direct Submission			
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
3 (bases 1 to 186403)			
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iltiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggiano,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,			
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
Direct Submission			
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
On Dec 1, 2001 this sequence version replaced gi:16974232.			
All repeats were identified using RepeatMasker:			
Smit, A.F.A. & Green, P. (1996-1997)			
http://ftp.genome.washington.edu/RM/RepeatMasker.html			
----- Genome Center			
Center: Whitehead Institute/ MIT Center for Genome Research			
Center code: WIBR			
Web site: http://www-seq.wi.mit.edu			
Contact: sequence_submissions@genome.wi.mit.edu			
----- Project Information			
Center project name: L2570			
Center clone name: 90_B_9			

FEATURES	Location/Qualifiers		
source	1. .186403		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="15"		
	/map="15"		
	/clone="RP11-90B9"		
	/clone_lib="RPCI-11 Human Male BAC"		
	2. .73		
repeat_region	/rpt_family="L1p"		
repeat_region	146. .262		
	/rpt_family="L1M4"		
repeat_region	443. .713		
	/rpt_family="L1M4"		
repeat_region	complement(806. .1103)		
	/rpt_family="MER20B"		
repeat_region	complement(1196. .1361)		
	/rpt_family="MER82"		
repeat_region	1383. .1530		
	/rpt_family="MIR"		
repeat_region	complement(1551. .1847)		
	/rpt_family="AluSq"		
repeat_region	1914. .1949		
	/rpt_family="(CA)n"		
repeat_region	complement(1950. .2031)		
	/rpt_family="MER82"		
repeat_region	complement(2033. .2097)		
	/rpt_family="MER20B"		
repeat_region	2098. .2401		
	/rpt_family="AluSp"		
repeat_region	complement(2402. .2609)		
	/rpt_family="MER20B"		
repeat_region	2739. .2850		
	/rpt_family="L1MB3"		
repeat_region	2851. .3147		
	/rpt_family="AluY"		
repeat_region	3148. .3182		
	/rpt_family="L1MB3"		
repeat_region	complement(3536. .3848)		
	/rpt_family="AluSq"		
unsure	3563		
	/note="probably T"		
repeat_region	4018. .4276		
	/rpt_family="AluJb"		
repeat_region	4582. .4748		
	/rpt_family="L1ME"		
repeat_region	4749. .5036		
	/rpt_family="AluJb"		
repeat_region	5037. .5229		
	/rpt_family="L1ME"		
repeat_region	5230. .5538		
	/rpt_family="AluJb"		
repeat_region	5539. .5711		
	/rpt_family="L1ME"		
repeat_region	complement(5958. .6237)		
	/rpt_family="AluSc"		

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-207G14, 2000 bp overlap; the clone sequenced to the right is RP11-92K14, 2000 bp overlap. Actual start of this clone is at base position 67479 of RP11-207G14; actual end is at base position 13505 of RP11-92K14.

Data from AC013732 was used to finish the clone, AC073088.

Polymorphisms have been identified between AC013732 AC073088.

FEATURES

source

Location/Qualifiers
1..55713
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-323P8"
/clone_lib="RPCI-11"
1..146

repeat_region

/rpt_family="L1"

repeat_region

152..263

/rpt_family="MaLR"

repeat_region

288..455

/rpt_family="MER1_type"

repeat_region

506..802

/rpt_family="Alu"

repeat_region

777..849

/rpt_family="A-rich"

repeat_region

804..1355

/rpt_family="L1"

repeat_region

2561..2602

/rpt_family="AT-rich"

repeat_region

2644..3075

/rpt_family="L1"

repeat_region

2683..2704

/rpt_family="AT-rich"

repeat_region

2903..2948

/rpt_family="AT-rich"

repeat_region

3272..4082

/rpt_family="L1"

repeat_region

4083..4134

/rpt_family="ERV1"

repeat_region

4135..4390

/rpt_family="ERV1"

repeat_region

4392..4463

/rpt_family="ERV1"

repeat_region

4464..4591

/rpt_family="ERV1"

repeat_region

4592..4767

/rpt_family="L1"

repeat_region

5639..5882

/rpt_family="MER1_type"

repeat_region

5883..6179

/rpt_family="Alu"

repeat_region

6157..6179

/rpt_family="(A)n"

repeat_region

6180..6425

/rpt_family="MER1_type"

repeat_region

6222..6265

/rpt_family="AT-rich"

repeat_region

6281..6312

/rpt_family="AT-rich"

repeat_region

6893..6944

/rpt_family="AT-rich"

repeat_region

/rpt_family="AT-rich"

repeat_region

7018..7296

/rpt_family="Alu"

repeat_region

7474..7760

/rpt_family="L1"

repeat_region

7805..7965

/rpt_family="L1"

repeat_region

7990..8247

/rpt_family="MaLR"

repeat_region

8248..8516

/rpt_family="Alu"

repeat_region

8517..8690

/rpt_family="MaLR"

repeat_region

9638..9752

/rpt_family="MIR"

repeat_region

10151..11635

/rpt_family="L1"

repeat_region

12015..12071

/rpt_family="(TG)n"

repeat_region

12690..13000

/rpt_family="Alu"

repeat_region

12971..13000

/rpt_family="(A)n"

repeat_region

13184..13266

/rpt_family="ERV1"

repeat_region

14373..14504

/rpt_family="L1"

repeat_region

15153..15436

/rpt_family="L1"

repeat_region

15680..16605

/rpt_family="L1"

repeat_region

16613..17007

/rpt_family="MaLR"

repeat_region

16755..16805

/rpt_family="(TC)n"

repeat_region

17008..17241

/rpt_family="L1"

repeat_region

17242..17543

/rpt_family="Alu"

repeat_region

17523..17543

/rpt_family="(A)n"

repeat_region

17544..17670

/rpt_family="L1"

repeat_region

17728..17759

/rpt_family="(TG)n"

repeat_region

17800..18266

/rpt_family="L1"

repeat_region

18267..18571

/rpt_family="Alu"

repeat_region

18572..19590

/rpt_family="L1"

repeat_region

19242..19340

/rpt_family="A-rich"

repeat_region

19594..19772

/rpt_family="L1"

repeat_region

19782..20196

/rpt_family="L1"

repeat_region

20222..20884

/rpt_family="L1"

repeat_region

20903..22061

/rpt_family="L1"

repeat_region

22442..22896

/note="similar to EST AW991366 (NID:98251411)"

repeat_region

22555..22662

/rpt_family="T-rich"

repeat_region

23176..25314

/rpt_family="L1"

repeat_region

25328..25489

/rpt_family="Alu"

repeat_region

25462..25497

Alignment Scores:

Pred. No.:

Score:

884

54.00

Length:

Matches:

55713

13


```

Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,F., Johnson,R., Jollivet,S., Joudah,S.,
Karissone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oraquye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivas,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,K.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 176203)
Worley,K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 176203)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 9, 2002 this sequence version replaced gi:17941058.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GAKJ
Center clone name: CH230-3J13
----- Summary Statistics
Sequencing vector: plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 133645 bases at least Q40
Consensus quality: 139597 bases at least Q30
Consensus quality: 143465 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 62 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

* be preserved.

```

1
1045: contig of 1045 bp in length
1146: gap of unknown length
1146: contig of 1575 bp in length
2721: gap of unknown length
2820: contig of 1078 bp in length
3898: gap of unknown length
3999: contig of 1210 bp in length
5209: gap of unknown length
5309: contig of 1250 bp in length
6559: gap of unknown length
7907: contig of 1249 bp in length
8007: gap of unknown length
9485: contig of 1478 bp in length
9585: gap of unknown length
10608: contig of 1023 bp in length
10709: gap of unknown length
12495: contig of 1787 bp in length
12595: gap of unknown length
13654: contig of 1059 bp in length
13754: gap of unknown length
15125: contig of 1371 bp in length
15225: gap of unknown length
16288: contig of 1063 bp in length
16388: gap of unknown length
18461: contig of 2073 bp in length
18561: gap of unknown length
19771: contig of 1210 bp in length
19871: gap of unknown length
21886: contig of 2015 bp in length
21986: gap of unknown length
23857: contig of 1871 bp in length
23957: gap of unknown length
25478: contig of 1521 bp in length
25578: gap of unknown length
27590: contig of 2012 bp in length
27690: gap of unknown length
29072: contig of 1382 bp in length
29172: gap of unknown length
31414: contig of 2242 bp in length
31514: gap of unknown length
33681: contig of 2167 bp in length
33781: gap of unknown length
35801: contig of 2020 bp in length
35901: gap of unknown length
37295: contig of 1394 bp in length
37395: gap of unknown length
39262: contig of 1867 bp in length
39362: gap of unknown length
41122: contig of 1760 bp in length
41222: gap of unknown length
42968: contig of 1746 bp in length
43068: gap of unknown length
44637: contig of 1569 bp in length
44737: gap of unknown length
46698: contig of 1961 bp in length
46798: gap of unknown length
48941: contig of 2143 bp in length
49041: gap of unknown length
51116: contig of 2075 bp in length
51216: gap of unknown length
51217: contig of 1909 bp in length
53126: gap of unknown length
53225: gap of unknown length
54938: contig of 1713 bp in length
55038: gap of unknown length
55039: contig of 2958 bp in length
57997: gap of unknown length
58096: contig of 1926 bp in length
60022: gap of unknown length
60122: gap of unknown length
63444: contig of 3222 bp in length
63445: gap of unknown length
65874: contig of 2430 bp in length
65974: gap of unknown length

```


* 1062 1161: gap of unknown length
 * 1162 2485: contig of 1324 bp in length
 * 2486 2585: gap of unknown length
 * 2586 3693: contig of 1108 bp in length
 * 3694 3794: gap of unknown length
 * 3794 5268: contig of 1475 bp in length
 * 5269 5368: gap of unknown length
 * 5369 6766: contig of 1398 bp in length
 * 6767 6866: gap of unknown length
 * 6867 8147: contig of 1281 bp in length
 * 8148 8248: gap of unknown length
 * 8249 9727: contig of 1480 bp in length
 * 9728 9828: gap of unknown length
 * 9829 11062: contig of 1235 bp in length
 * 11063 11162: gap of unknown length
 * 11163 12390: contig of 1228 bp in length
 * 12391 12490: gap of unknown length
 * 12491 14165: contig of 1675 bp in length
 * 14166 14265: gap of unknown length
 * 14266 15739: contig of 1474 bp in length
 * 15740 15839: gap of unknown length
 * 15840 18806: contig of 2967 bp in length
 * 18807 18906: gap of unknown length
 * 18907 22833: contig of 3927 bp in length
 * 22834 22933: gap of unknown length
 * 22934 27301: contig of 4368 bp in length
 * 27302 27401: gap of unknown length
 * 27402 34418: contig of 7017 bp in length
 * 34419 34518: gap of unknown length
 * 34519 39860: contig of 5342 bp in length
 * 39861 39960: gap of unknown length
 * 39961 47538: contig of 7578 bp in length
 * 47539 47638: gap of unknown length
 * 47639 53621: contig of 5983 bp in length
 * 53622 53721: gap of unknown length
 * 53722 72770: contig of 19049 bp in length
 * 72771 72870: gap of unknown length
 * 72871 88310: contig of 15440 bp in length
 * 88311 88411: gap of unknown length
 * 88411 106609: contig of 18199 bp in length
 * 106610 106709: gap of unknown length
 * 106710 130355: contig of 23646 bp in length
 * 130356 130456: gap of unknown length
 * 130456 174069: contig of 43614 bp in length
 * 174070 174169: gap of unknown length
 * 174170 231995: contig of 57826 bp in length.

FEATURES

Location/Qualifiers

1..231995
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /clone="CH230-216B10"

BASE COUNT 67304 a 43019 c 43334 g 65782 t 12556 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4,09e+03 Length: 231995
 Score: 54.00 Matches: 11
 Percent Similarity: 77.78% Conservative: 3
 Best Local Similarity: 61.11% Mismatches: 4
 Query Match: 51.92% Indels: 0
 DB: 2 Gaps: 0

US-09-833-017B-4 (1-21) x AC115352 (1-231995)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAla 18
 Db 226499 TCAGGCTCAACCCCTGTTCTTTCTCTTCATCGTGCATTACACAGTCA 226446

RESULT 13

WTVDI/c WTVDI 2613 bp ds-RNA linear VRL 03-AUG-1993
 LOCUS Wound tumor virus genome segment S5 encoding polypeptide P5,
 DEFINITION complete cds.

ACCESSION J03020
 VERSION J03020.1 GI:336169
 KEYWORDS polypeptide P5.
 SOURCE Wound tumor virus (inoculum RB) RNA, passed in *Agallia constricta* cell line AC20.
 ORGANISM Wound tumor virus
 REFERENCE 1 (bases 1 to 2613)
 Anzola, J. V., Xu, Z. K., Asamizu, T. and Nuss, D. L.
 Segment-specific inverted repeats found adjacent to conserved terminal sequences in wound tumor virus genome and defective interfering RNAs
 TITLE Proc. Natl. Acad. Sci. U.S.A. 84 (23), 8301-8305 (1987)
 JOURNAL 88068582
 MEDLINE 3479793
 COMMENT Draft entry and printed copy of sequence for [1] kindly provided by D.L. Nuss, 10/13/87.
 FEATURES
 source location/Qualifiers
 1..2613
 /organism="Wound tumor virus"
 /db_xref="taxon:10987"
 26..2440
 /note="polypeptide P5"
 /codon_start=1
 /protein_id="AA048499.1"
 /db_xref="GI:336170"
 /translation="MAIDSYICIPNFQSQTIDNRTIVNIFOSCKYRSQSLVCFLNKDSAA
 DRFSNMRQSGTITFIHAEDGEISEQLHSTFRSVSTMLCGMLFVIFAPRNVIS
 SETGKAITWAFRGSFIELDRHQEQALHDLFQFRLSLPLVNVKMGWAYGPTSFA
 ELISLKNKTSWRYVIDYSMTFSALGVFASHMDECSFANKOINVIGVNPYVWAG
 LRHGVTRTFETSTPDEGYPKILPLRTGTNVLKKYKVQHDQPKLLCDDSYME
 ALSRNILYIGVYPATHLLDNLKGRVAVDPKINAAEAETLKQRTSIDLYLSARFE
 FNAQSTRDIVLYKFSGVFFSIDDSWVEGTEDEYKQELKQSTFEQLVWNGSTSKURV
 SMISMKNRTRDKVCKRRLALLPQYGGSLREURAFHVNGAAEVNKKSEVNSYMDK
 FTSLSISEQSGQKFMHMLTNTYDALKTKTGDKATIIAYSLSNAINKKRVKFLS
 DAAKSETLIIFGAPLNVRVFMKISGVILGSDVTISNDLITFNKASGVKWDYQTS
 ELIKSSMIEITIEQMLNCRVSSYNGVGFANSYINDMFSWYVPEMLFEKYSIQDIRL
 SPVALVKCFTTSIRNLCTYVALRGSEFVEKVLITNNVLNSSLVLTGTHSTPRL
 LSNFEVPSAGVLKFKAGDDVNTSGHLLSLVIAAFVASPTLLWATMKRMHTPTNLP
 KNLDKLLFFDNKIKNGMKHRSREVVLAAMIVENVAHILNGRHSIEIIOITQVI
 YERFNA"
 BASE COUNT 825 a 428 c 545 g 815 t
 ORIGIN Unreported.
 Alignment Scores:
 Pred. No.: 48.4 Length: 2613
 Score: 53.00 Matches: 10
 Percent Similarity: 68.42% Conservative: 3
 Best Local Similarity: 52.63% Mismatches: 6
 Query Match: 50.96% Indels: 0
 DB: 14 Gaps: 0
 US-09-833-017B-4 (1-21) x WTVDI (1-2613)
 QY 2 GlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 816 GGATCATGCTGACGTATTTTACCTTCTTTAACAGGACATTTCCAGTAAGCCTAGGT 760
 RESULT 14
 AL512364 133226 bp DNA linear HTG 10-JUL-2001
 LOCUS Homo sapiens chromosome 6 clone RP11-574F19, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 25 unordered pieces.
 ACCESSION AL512364
 VERSION AL512364.3 GI:11995201
 KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 133226)
 AUTHORS Burton, J.
 TITLE Direct Submission

JOURNAL

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Dec 28, 2000 this sequence version replaced gi:11995149.

COMMENT

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA574F19
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: Plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 120707 bases at least Q40
Consensus quality: 125263 bases at least Q30
Consensus quality: 127799 bases at least Q20
Insert size: 130826; sum-of-contigs
Insert size: 158221; 10.3% error; agarose-fp
Quality coverage: 3.07x in Q20 bases; sum-of-contigs Quality
coverage: 2.92x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

* 1 4495: contig of 4495 bp in length
* 4496 4595: gap of 100 bp
* 4596 17538: contig of 12943 bp in length
* 17539 17638: gap of 100 bp
* 17639 21198: contig of 3560 bp in length
* 21199 21298: gap of 100 bp
* 21299 24416: contig of 3118 bp in length
* 24417 24516: gap of 100 bp
* 24517 27845: contig of 3329 bp in length
* 27846 27945: gap of 100 bp
* 27946 35081: contig of 7136 bp in length
* 35082 35181: gap of 100 bp
* 35182 37225: contig of 2044 bp in length
* 37226 37325: gap of 100 bp
* 37326 39611: contig of 2286 bp in length
* 39612 39711: gap of 100 bp
* 39712 42865: contig of 3154 bp in length
* 42866 42965: gap of 100 bp
* 42966 49369: contig of 6404 bp in length
* 49370 49469: gap of 100 bp
* 49470 52300: contig of 2731 bp in length
* 52201 52300: gap of 100 bp
* 52301 55073: contig of 2773 bp in length
* 55074 55173: gap of 100 bp
* 55174 62527: contig of 7354 bp in length
* 62528 62627: gap of 100 bp
* 62628 66039: contig of 3412 bp in length
* 66040 66139: gap of 100 bp
* 66140 69939: contig of 3800 bp in length
* 69940 70039: gap of 100 bp
* 70040 78998: contig of 8959 bp in length
* 78999 79098: gap of 100 bp
* 79099 84516: contig of 5418 bp in length
* 84517 84616: gap of 100 bp
* 84617 93119: contig of 8503 bp in length
* 93120 93219: gap of 100 bp
* 93220 95763: contig of 2544 bp in length
* 95764 95863: gap of 100 bp
* 95864 109619: contig of 13756 bp in length
* 109620 109719: gap of 100 bp
* 109720 111843: contig of 2124 bp in length
* 111844 111943: gap of 100 bp

* 111944 114312: contig of 2369 bp in length
* 114313 114412: gap of 100 bp
* 114413 126023: contig of 11611 bp in length
* 126024 126123: gap of 100 bp
* 126124 130297: contig of 4174 bp in length
* 130298 130397: gap of 100 bp
* 130398 133226: contig of 2829 bp in length.
FEATURES
Location/Qualifiers
1. .133226
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-574F19"
/clone_lib="RPC1-11.2"
1. .4495
/note="assembly_fragment:00744"
fragment_chain:1"
4596. .17538
/note="assembly_fragment:00397"
fragment_chain:1"
17639. .21198
/note="assembly_fragment:00155"
fragment_chain:1"
21299. .24416
/note="assembly_fragment:00059"
fragment_chain:2"
24517. .27845
/note="assembly_fragment:00653"
fragment_chain:2"
27946. .35081
/note="assembly_fragment:00142"
fragment_chain:3"
35182. .37225
/note="assembly_fragment:01217"
fragment_chain:3"
37326. .39611
/note="assembly_fragment:00287"
fragment_chain:4"
39712. .42865
/note="assembly_fragment:01572"
fragment_chain:4"
42966. .49369
/note="assembly_fragment:00349"
fragment_chain:5"
49470. .52200
/note="assembly_fragment:00572"
fragment_chain:5"
52301. .55073
/note="assembly_fragment:00399"
fragment_chain:6"
55174. .62527
/note="assembly_fragment:01761"
fragment_chain:6"
62628. .66039
/note="assembly_fragment:00508"
fragment_chain:7"
66140. .69939
/note="assembly_fragment:01732"
fragment_chain:7"
70040. .78998
/note="assembly_fragment:01083"
fragment_chain:8"
79099. .84516
/note="assembly_fragment:00253"
fragment_chain:8"
84617. .93119
/note="assembly_fragment:01719"
fragment_chain:9"
93220. .95763
/note="assembly_fragment:01419"
fragment_chain:9"
95864. .109619
/note="assembly_fragment:00001"

misc_feature 109720..111843
 /note="assembly_fragment:00525"
 misc_feature 111944..114312
 /note="assembly_fragment:00785"
 misc_feature 114413..126023
 /note="assembly_fragment:01103"
 misc_feature 126124..130297
 /note="assembly_fragment:01207"
 misc_feature 130398..133226
 /note="assembly_fragment:01360"
 BASE COUNT 38629 a 24563 c 25616 g 41987 t 2431 others
 ORIGIN

Alignment Scores: 3.3e+03 Length: 133226
 Pred. No.: 53.00 Matches: 11
 Score: 75.00% Conservative: 1
 Percent Similarity: 68.75% Mismatches: 4
 Best Local Similarity: 50.96% Indels: 0
 Query Match: 2 Gaps: 0
 DB: 2

US-09-833-017B-4 (1-21) x AL512364 (1-133226)

Oy 2 GlySerLeuSerThrPheArgLeuPheAsnArgSerPheThrGln 17
 Db 46992 GGTTGGTGAAGCACGTTTGGGAGGCTTTTAAACAAAGTTTCACACAG 47039

RESULT 15

AL513497/c

LOCUS

Human DNA sequence from clone RP4-669K10 on chromosome 1, complete

sequence.

AL513497.21 GI:22204196

HTG.

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 135820)

White,S.

Direct Submission

Submitted (07-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,

Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Aug 11, 2002 this sequence version replaced gi:21530913.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em., EMBL; Sw.,

SWISSPROT; Tr., TrEMBL; Wp., WormPep; Information on the WormPep

database can be found at

http://www.sanger.ac.uk/projects/c-elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1

RP4-669K10 is from the library RPI-4 constructed by the group of

Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pcypac2

----- Genome Center

Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: humquery@sanger.ac.uk

FEATURES

Source

Location/Qualifiers
 1..135820
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP4-669K10"
 /clone_lib="RPI-4"

BASE COUNT 35410 a 32344 c 33048 g 35018 t
 ORIGIN

Alignment Scores: 3.37e+03 Length: 135820
 Pred. No.: 53.00 Matches: 11
 Score: 77.78% Conservative: 3
 Percent Similarity: 61.11% Mismatches: 4
 Best Local Similarity: 50.96% Indels: 0
 Query Match: 9 Gaps: 0
 DB: 9

US-09-833-017B-4 (1-21) x AL513497 (1-135820)

Oy 1 SerGlySerLeuSerThrPheArgLeuPheAsnArgSerPheThrGlnAla 18

Db 40215 AGCGGCTCAGTAAGCCTTTTGTGAGATGGAGTCGCGCTTCTTCCACCGCT 40162

Search completed: November 8, 2002, 22:57:36

Job time : 837.284 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 20:29:27 ; Search time 104.687 Seconds
(without alignments)
451.748 Million cell updates/sec

Title: US-09-833-017b-4

Perfect score: 104

Sequence: 1 SGSLSTFPFLNRSFTQALGK 21

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09833017/runat_05112002_105348_4824/app.query.fasta_1.398
-DB=N_Geneseq_101002 -QFMT=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPTYPE=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017.ecgn_1.1.125_erunat_05112002_105348_4824 -NCPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001C.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	104	100.0	63	24	AAD32790	Streptococcus muta
2	104	100.0	63	24	AAD32884	Streptococcus muta
3	104	100.0	141	24	AAD32791	Streptococcus muta
4	104	100.0	141	24	AAD32898	Streptococcus muta
5	104	100.0	2557	24	AAD32800	Streptococcus muta
6	104	100.0	2557	24	AAD32893	Streptococcus muta
7	52	50.0	10475	9	AA80616	Genomic clone enco
8	52	50.0	10475	9	AA80643	Sequence of human
9	52	50.0	10476	24	ABL67064	Myroid cancer rel
10	52	50.0	10476	24	ABL67518	Myroid cancer rel
11	50	48.1	2017	22	AAS56980	C. trachomatis DNA
12	50	48.1	2932	23	ABL04050	Drosophila melanog
13	50	48.1	3596	23	ABL03962	Drosophila melanog
14	49	47.1	694	19	AAV27411	Streptococcus pneu
15	49	47.1	694	24	ABQ84879	S. pneumoniae Sp10
16	49	47.1	1017	21	AAC41522	Arabidopsis thalia
17	49	47.1	1099	21	AAC32612	Arabidopsis thalia
18	49	47.1	1245	21	AAC35111	Arabidopsis thalia
19	49	47.1	2753	13	AAQ30801	p18 plasmid. Syn
20	49	47.1	2784	23	ABL08342	Drosophila melanog
21	49	47.1	3319	13	AAQ30802	RI gene. Syntheti
22	49	47.1	3671	24	ABK84195	Human cDNA differe
23	49	47.1	3690	22	AAS31177	Human diagnostic a
24	49	47.1	4276	23	ABL11534	Drosophila melanog
25	49	47.1	6004	19	AAV52160	Streptococcus pneu
26	48.5	46.6	370	23	ABV37598	Human prostate exp
27	48	46.2	369	22	AAI98602	Human excretory re
28	48	46.2	369	22	AAI62998	Human kidney relat
29	48	46.2	885	24	ABQ46476	Oligonucleotide fo
30	48	46.2	885	24	ABQ46477	Oligonucleotide fo
31	48	46.2	1167	22	AAS2612	S. epidermidis ope
32	48	46.2	1317	22	ABN90839	Staphylococcus epi
33	48	46.2	3191	22	AAH54281	S. epidermidis gen
34	48	46.2	3285	22	AAH54084	S. epidermidis gen
35	48	46.2	3391	22	AAH54393	S. epidermidis gen
36	48	46.2	3882	23	ABL20364	Drosophila melanog
37	48	46.2	319608	21	AAH51601	Human chromosome 1
38	48	46.2	319608	22	AAH509301	Human chizophreni
39	47.5	45.7	565	24	ABN65805	Human cancer relat
40	47.5	45.7	3679	24	ABA93705	Human nucleic acid
41	47.5	45.7	3776	23	ABL22430	Drosophila melanog
42	47.5	45.2	870	24	ABN98796	Arabidopsis thalia
43	47	45.2	902	24	ABK30867	Plant dwarfing/stu
44	47	45.2	902	24	ABK30952	Plant dwarfing/stu
45	47	45.2	945	24	ABL49401	Sequence #3 used t

ALIGNMENTS

RESULT 1

AAD32790

ID AAD32790 standard; DNA; 63 Bp.

XX

AC AAD32790;

XX 01-JUL-2002 (first entry)

DT

XX Streptococcus mutans competence signal peptide (CSP) DNA.

DE Competence signal peptide; CSP; histidine kinase; HK; therapy; carles;
KW endocarditis; immunopurification; antibacterial; antiinflammatory;
KW genetic competence assay; vaccine; gene; ds.
XX Streptococcus mutans.

OS

XX Key Location/Qualifiers

FF CDS 1..63

FT

```

FT      /*tag= a
FT      /product= "S. mutans competence signal peptide (CSP)"
FT      /note= "CDS does not include start and stop codon"
FT      /partial
XX      CA2302861-A1.
XX      10-OCT-2001.
XX
XX      10-APR-2000; 2000CA-2302861.
XX
XX      10-APR-2000; 2000CA-2302861.
XX      (LAUP/) LAU P C Y.
XX      (CVIT/) CVITKOVITCH D G.
XX      (LIYH/) LI Y H.
XX
XX      Lau PCY, Cvitkovitch DG, Li YH;
XX      WPI; 2002-242151/30.
XX      P-PSDB; AAE20493.
XX
XX      Novel compound that inhibits binding of competence signal peptide of
XX      Streptococcus mutans to S. mutans histidine kinase, useful for treating
XX      or preventing caries or endocarditis -
XX      Claim 8; Fig 2; 54pp; English.
XX
XX      The invention relates to compound that competitively inhibits binding
XX      of competence signal peptide (CSP) to Streptococcus mutans histidine
XX      kinase (HK). Compounds of the invention are useful for treating or
XX      prophylaxis of caries or endocarditis. Vector comprising nucleic acid
XX      encoding S. mutans CSP is useful for producing recombinant host cell
XX      capable of expressing it. The recombinant host cell produced by the
XX      method is useful for expressing peptide in culture. S. mutans CSP DNA
XX      is useful for identifying nucleic acid molecules encoding CSP activated
XX      peptide. It is also useful as probes and in assays to identify
XX      antagonists or inhibitors of the peptides produced by the nucleic acid
XX      molecules. It is also useful for preparing vaccines for preventing or
XX      treating the above mentioned conditions. Antibodies against CSP activity
XX      are also useful for preventing caries. The antibodies are also useful
XX      for screening organisms or tissues containing CSP peptide or CSP-like
XX      peptide and for immunopurifying the peptides. The CSP nucleic acid
XX      molecules are useful in assays for genetic competence. The present
XX      sequence is S. mutans CSP DNA.
XX      Sequence 63 BP; 20 A; 12 C; 13 G; 18 T; 0 other;

Alignment Scores:
Pred. No.:      2.65e-10      Length:      63
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:             24      Gaps:
US-09-833-017B-4 (1-21) x AAD32790 (1-63)

QY      1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
Db      1 AGCGGAAGCCTATCAACATTTTCCGCGCTGTTTAAACAGAAGTTTACACAAAGCTTTGGCA 60

QY      21 Lys 21
Db      61 AAA 63

RESULT 2
AAD32884
ID      AAD32884 standard; DNA; 63 BP.
XX
XX      AAD32884;
XX
XX      01-JUL-2002 (first entry)

```

```

XX      Streptococcus mutans competence signal peptide (CSP) DNA.
XX      Competence signal peptide; CSP; histidine kinase; HK; prophylaxis;
XX      therapy; caries; endocarditis; microbial biofilm; infection; gene;
XX      vaccine; antibacterial; ds.
XX      Streptococcus mutans.
XX
XX      Key      Location/Qualifiers
XX      CDS      1..63
XX              /tag= a
XX              /product= "S. mutans competence signal peptide"
XX              /note= "CDS does not include start and stop codon"
XX              /partial
XX
XX      CA2332733-A1.
XX
XX      10-OCT-2001.
XX
XX      20-FEB-2001; 2001CA-2332733.
XX
XX      10-APR-2000; 2000CA-2302861.
XX      (LIYH/) LI Y H.
XX      (CVIT/) CVITKOVITCH D G.
XX      (LAUP/) LAU P C.
XX
XX      Li YH, Cvitkovitch DG, Lau PC;
XX      WPI; 2002-242173/30.
XX      P-PSDB; AAE20584.
XX
XX      Novel compound that competitively inhibits binding of competence signal
XX      peptide to Streptococcus mutans histidine kinase, useful in treatment
XX      or prophylaxis of caries or endocarditis -
XX      Claim 8; Fig 2; 82pp; English.
XX
XX      The invention relates to a compound that competitively inhibits the
XX      binding of competence signal peptide (CSP) to Streptococcus mutans
XX      histidine kinase (HK). Compounds of the invention are useful in
XX      medical treatment or prophylaxis of caries or endocarditis. They are
XX      useful for inhibiting or disrupting microbial biofilms involved in
XX      infections in man and animals and in biofouling of surfaces susceptible
XX      to microbial accumulation. They are also useful for treatment or
XX      prophylaxis of a disease, disorder or abnormal physical state caused
XX      by S. mutans. Pharmaceutical composition containing the compounds of
XX      the invention is useful for treating diseases caused by streptococcal
XX      infections. Polynucleotides encoding S. mutans CSP are useful as probes
XX      or in assays to identify antagonists or inhibitors of CSP peptides.
XX      They are also used for preparing vaccines. Polypeptides of the invention
XX      are useful for preparing antibodies, for vitro analysis of HK, CSP or
XX      RR activity or structure, and in assays for the identification and
XX      developments of compounds to inhibit and/or enhance polypeptide or
XX      peptide function directly. Antibodies of the invention are useful for
XX      providing protection against caries, to screen organisms or tissues
XX      containing CSP peptide or CSP-like peptides, for immuno-purification of
XX      CSP or CSP-like peptides from crude extracts, and to detect CSP or a
XX      similar peptide. The present sequence is S. mutans CSP DNA.
XX      Sequence 63 BP; 20 A; 12 C; 13 G; 18 T; 0 other;

Alignment Scores:
Pred. No.:      2.65e-10      Length:      63
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:             24      Gaps:
US-09-833-017B-4 (1-21) x AAD32884 (1-63)

```

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 1 AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAAAGAGAAGTTTACACAAGCTTTGGGA 60
 Qy 21 Lys 21
 Db 61 AAA 63
 RESULT 3
 AAD32791
 ID AAD32791 standard; DNA; 141 BP.
 XX
 AC AAD32791;
 XX
 DT 01-JUL-2002 (first entry)
 XX Streptococcus mutans comC gene.
 DE
 XX Competence signal peptide; CSP; histidine kinase; HK; therapy; caries;
 KW endocarditis; immunopurification; antibacterial; antiinflammatory;
 KW genetic competence assay; vaccine; gene; comC gene; ds.
 XX
 OS Streptococcus mutans.
 XX
 XX Key Location/Qualifiers
 FH CDS 1..141
 FT /*tag= a
 FT /product= "S. mutans ComC protein"
 XX CA2302861-A1.
 XX 10-OCT-2001.
 XX 10-APR-2000; 2000CA-2302861.
 XX 10-APR-2000; 2000CA-2302861.
 XX (LAUP/) LAU P C Y.
 PA (CVIT/) CVITKOVITCH D G.
 PA (LIYH/) LI Y H.
 XX
 PI Lau PCY, Cvitkovitch DG, Li YH;
 XX WPI; 2002-242151/30.
 DR P-PSDB; AAE20494.
 XX Novel compound that inhibits binding of competence signal peptide of
 PT Streptococcus mutans to S. mutans histidine kinase, useful for treating
 PT or preventing caries or endocarditis
 XX Disclosure; Fig 2; 54pp; English.
 XX The invention relates to compound that competitively inhibits binding
 CC of competence signal peptide (CSP) to Streptococcus mutans histidine
 CC kinase (HK). Compounds of the invention are useful for treating or
 CC prophylaxis of caries or endocarditis. Vector comprising nucleic acid
 CC encoding S. mutans CSP is useful for producing recombinant host cell
 CC capable of expressing it. The recombinant host cell produced by the
 CC method is useful for expressing peptide in culture. S. mutans CSP DNA
 CC is useful for identifying nucleic acid molecules encoding CSP activated
 CC peptide. It is also useful as probes and in assays to identify
 CC antagonists or inhibitors of the peptides produced by the nucleic acid
 CC molecules. It is also useful for preparing vaccines for preventing or
 CC treating the above mentioned conditions. Antibodies against CSP activity
 CC are also useful for preventing caries. The antibodies are also useful
 CC for screening organisms or tissues containing CSP peptide or CSP-like
 CC peptide and for immunopurifying the peptides. The CSP nucleic acid
 CC sequence is S. mutans comC CSP precursor gene.
 XX Sequence 141 BP; 58 A; 19 C; 25 G; 39 T; 0 other;
 SQ Alignment Scores:

Pred. No.: 7.13e-10 Length: 141
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-833-017B-4 (1-21) x AAD32791 (1-141)
 Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 Db 76 AGCGGAAGCCTATCAACATTTTCCGGCTGTTTAAAGAGAAGTTTACACAAGCTTTGGGA 135
 Qy 21 Lys 21
 Db 136 AAA 138
 RESULT 4
 AAD32898
 ID AAD32898 standard; DNA; 141 BP.
 XX
 AC AAD32898;
 XX
 DT 01-JUL-2002 (first entry)
 XX Streptococcus mutans ComC gene.
 DE
 XX Competence signal peptide; CSP; histidine kinase; HK; prophylaxis;
 KW therapy; caries; endocarditis; microbial biofilm; infection; gene;
 KW vaccine; antibacterial; ComC gene; ds.
 XX
 OS Streptococcus mutans.
 XX
 XX Key Location/Qualifiers
 FH CDS 1..141
 FT /*tag= a
 FT /product= "S. mutans ComC protein"
 XX CA232733-A1.
 XX 10-OCT-2001.
 XX 20-FEB-2001; 2001CA-2332733.
 XX 10-APR-2000; 2000CA-2302861.
 XX (LIYH/) LI Y H.
 PA (CVIT/) CVITKOVITCH D G.
 PA (LAUP/) LAU P C.
 XX
 PI Li YH, Cvitkovitch DG, Lau PC;
 XX WPI; 2002-242173/30.
 DR P-PSDB; AAE20594.
 XX Novel compound that competitively inhibits binding of competence signal
 PT peptide to Streptococcus mutans histidine kinase, useful in treatment
 PT or prophylaxis of caries or endocarditis
 XX Disclosure; Fig 2; 82pp; English.
 XX The invention relates to a compound that competitively inhibits the
 CC binding of competence signal peptide (CSP) to Streptococcus mutans
 CC histidine kinase (HK). Compounds of the invention are useful in
 CC medical treatment or prophylaxis of caries or endocarditis. They are
 CC useful for inhibiting or disrupting microbial biofilms involved in
 CC infections in man and animals and in biofouling of surfaces susceptible
 CC to microbial accumulation. They are also useful for treatment or
 CC prophylaxis of a disease, disorder or abnormal physical state caused
 CC by S. mutans. Pharmaceutical composition containing the compounds of
 CC the invention is useful for treating diseases caused by streptococcal
 CC infections. Polynucleotides encoding S. mutans CSP are useful as probes
 CC or in assays to identify antagonists or inhibitors of CSP peptides.

CC They are also used for preparing vaccines. Polypeptides of the invention
 CC are useful for preparing antibodies, for vitro analysis of HK, CSP or
 CC RR activity or structure, and in assays for the identification and
 CC development of compounds to inhibit and/or enhance polypeptide or
 CC peptide function directly. Antibodies of the invention are useful for
 CC providing protection against caries, to screen organisms or tissues
 CC containing CSP peptide or CSP-like peptides, for immuno-purification of
 CC CSP or CSP-like peptides from crude extracts, and to detect CSP or a
 CC similar peptide. The present sequence is S. mutans comC gene encoding a
 CC CSP precursor protein.

XX SQ Sequence 141 BP; 58 A; 19 C; 25 G; 39 T; 0 other;

Alignment Scores:
 Pred. No.: 7.13e-10 Length: 141
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-833-017B-4 (1-21) x AAD32898 (1-141)

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 |||||
 Db 76 AGCGGAAGCCTATCAACATTTTCGCGCTGTTTACAGAGATTTTACACAAAGCTTTGGGA 135

Qy 21 Lys 21

Db 136 AAA 138

RESULT 5

AAD32800
 ID AAD32800 standard; DNA; 2557 BP.

XX AC AAD32800;

XX DT 01-JUL-2002 (first entry)

XX DE Streptococcus mutans comCDE gene local region.

XX KW Competence signal peptide; CSP; histidine kinase; HK; therapy; caries;
 KW endocarditis; immunopurification; antibacterial; antiinflammatory;
 KW genetic competence assay; vaccine; comCDE gene; gene; ds.

XX OS Streptococcus mutans.

XX FH Location/Qualifiers
 FT complement (25..168)
 FT CDS /*tag= a
 FT /product= "Protein #1 encoded by S. mutans comCDE gene
 FT local region"
 FT complement (452..499)
 FT CDS /*tag= b
 FT /product= "Peptide #1 encoded by S. mutans comCDE gene
 FT local region"
 FT /note= "CDS does not include start and stop codon"
 FT partial
 FT 648..758

FT CDS /*tag= c
 FT /product= "Peptide #2 encoded by S. mutans comCDE gene
 FT local region"
 FT 953..1081
 FT CDS /*tag= d
 FT /product= "Protein #2 encoded by S. mutans comCDE gene
 FT local region"
 FT complement (1366..1449)
 FT /*tag= e

FT /product= "Peptide #3 encoded by S. mutans comCDE gene
 FT local region"
 FT 1855..1959
 FT CDS /*tag= f
 FT /product= "Peptide #4 encoded by S. mutans comCDE gene

FT CDS local region"
 FT complement (1896..2072)
 FT /*tag= g
 FT /product= "Protein #3 encoded by S. mutans comCDE gene
 FT local region"
 FT 2182..2424
 FT CDS /*tag= h
 FT /product= "Protein #4 encoded by S. mutans comCDE gene
 FT local region"
 FT 2384..2488
 FT CDS /*tag= i
 FT /product= "Peptide #5 encoded by S. mutans comCDE gene
 FT local region"

XX CA2302861-A1.

XX PD 10-OCT-2001.

XX PF 10-APR-2000; 2000CA-2302861.

XX PR 10-APR-2000; 2000CA-2302861.

XX PA (LAUF/) LAU P C Y.

XX PA (CVIT/) CVITKOVITCH D G.

XX PA (LIYH/) LI Y H.

XX PI Lau PCY, Cvitkovitch DG, Li YH;

XX XX WPI; 2002-242151/30.

DR P-PSDB; AAE20607, AAE20608, AAE20609, AAE20610, AAE20611, AAE20612,

DR AAE20613, AAE20614, AAE20615.

XX Novel compound that inhibits binding of competence signal peptide of
 PT Streptococcus mutans to S. mutans histidine kinase, useful for treating
 PT or preventing caries or endocarditis

XX PS Disclosure; Fig 9; 54pp; English.

XX The invention relates to compound that competitively inhibits binding
 CC of competence signal peptide (CSP) to Streptococcus mutans histidine
 CC kinase (HK). Compounds of the invention are useful for treating or
 CC prophylaxis of caries or endocarditis. Vector comprising nucleic acid
 CC encoding S. mutans CSP is useful for producing recombinant host cell
 CC capable of expressing it. The recombinant host cell produced by the
 CC method is useful for expressing peptide in culture. S. mutans CSP DNA
 CC is useful for identifying nucleic acid molecules encoding CSP activated
 CC peptide. It is also useful as probes and in assays to identify
 CC antagonists or inhibitors of the peptides produced by the nucleic acid
 CC molecules. It is also useful for preparing vaccines for preventing or
 CC treating the above mentioned conditions. Antibodies against CSP activity
 CC are also useful for preventing caries. The antibodies are also useful
 CC for screening organisms or tissues containing CSP peptide or CSP-like
 CC peptide and for immunopurifying the peptides. The CSP nucleic acid
 CC molecules are useful in assays for genetic competence. The present
 CC sequence is S. mutans comCDE gene local region.

XX SQ Sequence 2557 BP; 891 A; 415 C; 426 G; 825 T; 0 other;

Alignment Scores:

Pred. No.: 2.5e-08 Length: 2557
 Score: 104.00 Matches: 21
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0

US-09-833-017B-4 (1-21) x AAD32800 (1-2557)

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
 |||||

Db 176 AGCGGAAGCCTATCAACATTTTCGCGCTGTTTACAGAGATTTTACACAAAGCTTTGGGA 235

Qy 21 Lys 21

```

Db      236 AAA 238
|||||
RESULT 6
AAD32893
ID      AAD32893 standard; DNA; 2557 BP.
XX
AC      AAD32893;
XX
DT      01-JUL-2002 (first entry)
XX
DE      Streptococcus mutans comCDE gene local region.
XX
KW      Competence signal peptide; CSP; histidine kinase; HK; prophylaxis;
KW      therapy; caries; endocarditis; microbial biofilm; infection; gene;
KW      vaccine; antibacterial; comCDE gene local region; ds.
XX
OS      Streptococcus mutans.
XX
FH      Key
CDS     complement (25...168)
        /tag= a
        /product= "Protein #1 encoded by S. mutans ComCDE
        gene local region"
        101...241
        /tag= b
        /product= "S. mutans ComC protein"
        complement (383...1708)
        /tag= c
        /product= "S. mutans ComD protein"
        complement (452...499)
        /tag= d
        /product= "Peptide #1 encoded by S. mutans ComCDE
        gene local region"
        /note= "CDS does not include start and stop codon"
        /partial
        648...758
        /tag= e
        /product= "Peptide #2 encoded by S. mutans ComCDE
        gene local region"
        953...1081
        /tag= f
        /product= "Protein #2 encoded by S. mutans ComCDE
        gene local region"
        complement (1366...1449)
        /tag= g
        /product= "Peptide #3 encoded by S. mutans ComCDE
        gene local region"
        complement (1705...2457)
        /tag= h
        /product= "S. mutans ComE protein"
        1855...1959
        /tag= i
        /product= "Peptide #4 encoded by S. mutans ComCDE
        gene local region"
        complement (1896...2072)
        /tag= j
        /product= "Protein #3 encoded by S. mutans ComCDE
        gene local region"
        2182...2424
        /tag= k
        /product= "Protein #4 encoded by S. mutans ComCDE
        gene local region"
        2384...2488
        /tag= l
        /product= "Peptide #5 encoded by S. mutans ComCDE
        gene local region"
XX
CA232733-AL.
XX
PD      10-OCT-2001.
XX
PF      20-FEB-2001; 2001CA-2332733.

```

```

XX
PR      10-APR-2000; 2000CA-2302861.
XX
PA      (LIYH/) LI Y H.
PA      (CVIT/) CVITKOVITCH D G.
PA      (LAUP/) LAU P C.
XX
PI      LI YH, Cvitkovitch DG, Lau PC;
XX
WI      MPI; 2002-242173/30.
DR      P-PSDB; AAE20617, AAE20618, AAE20619, AAE20620, AAE20621, AAE20622,
DR      AAE20623, AAE20624, AAE20625, AAE20594, AAE20585, AAE20586.
XX
PT      Novel compound that competitively inhibits binding of competence signal
PT      peptide to Streptococcus mutans histidine kinase, useful in treatment
PT      or prophylaxis of caries or endocarditis.
XX
PS      Disclosure; Fig 9; 82pp; English.
XX
CC      The invention relates to a compound that competitively inhibits the
CC      binding of competence signal peptide (CSP) to Streptococcus mutans
CC      histidine kinase (HK). Compounds of the invention are useful in
CC      medical treatment or prophylaxis of caries or endocarditis. They are
CC      useful for inhibiting or disrupting microbial biofilms involved in
CC      infections in man and animals and in biofouling of surfaces susceptible
CC      to microbial accumulation. They are also useful for treatment or
CC      prophylaxis of a disease, disorder or abnormal physical state caused
CC      by S. mutans. Pharmaceutical composition containing the compounds of
CC      the invention is useful for treating diseases caused by streptococcal
CC      infections. Polynucleotides encoding S. mutans CSP are useful as probes
CC      or in assays to identify antagonists or inhibitors of CSP peptides.
CC      They are also used for preparing vaccines. Polypeptides of the invention
CC      are useful for preparing antibodies, for vitro analysis of HK, CSP or
CC      HK activity or structure, and in assays for the identification and
CC      development of compounds to inhibit and/or enhance polypeptide or
CC      peptide function directly. Antibodies of the invention are useful for
CC      providing protection against caries, to screen organisms or tissues
CC      containing CSP peptide or CSP-like peptides, for immuno-purification of
CC      CSP or CSP-like peptides from crude extracts, and to detect CSP or a
CC      similar peptide. The present sequence is S. mutans comCDE gene local
CC      region.
XX
SQ      Sequence 2557 BP; 891 A; 415 C; 436 G; 825 T; 0 other;

Alignment Scores:
Pred. No.:      2.5e-08      Length:      2557
Score:          104.00      Matches:      21
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              24          Gaps:         0

US-09-833-017B-4 (1-21) x AAD32893 (1-2557)

Qy      1 SerGlySerLeuSerThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
        |||||
Db      176 AGCGGAGCGCTATCACATTTTCCGCGTGTTCACAGAGTTTACACAGCTTTGGGA 235

Qy      21 Lys 21
        |||
Db      236 AAA 238

RESULT 7
AAD30616/c
ID      AAN80616 standard; DNA; 10475 BP.
XX
AC      AAN80616;
XX
DT      03-OCT-2002 (updated)
DT      21-NOV-1990 (first entry)
XX
DE      Genomic clone encoding SAP(Phe).
XX

```


PI Whitsett JA, Fox JL, Pilot-Matias TJ, Meuth JL, Sarin VK;
 XX WPI; 1988-133244/19.
 DR P-PSDB; AAP80193, AAP80194, AAP80195, 081070, AAP81074, AAP81075,
 DR AAP81076, AAP81077, AAP81078, AAP81079.
 XX Pulmonary hydrophobic surfactant-associated proteins -
 PT used with lipid(s) to treat and prevent hyaline membrane disease
 PT and similar syndromes
 XX Example; Fig 8B-1 - 8B-24; 144pp; English.
 XX The sequence is derived from human embryonic kidney cell genomic
 CC library. SAP (Val) and SAP (Phe), when combined with lipids, have
 CC significant pulmonary biophysical surfactant activity that may be
 CC utilised to treat and prevent hyaline membrane disease (HMD) and
 CC other syndromes associated with lack or insufficient amts. of natural
 CC pulmonary surfactant material. Antibodies and antisera may also be
 CC made which are directed against SAP (Val) or SAP (Phe).
 CC SAP (Val) and SAP (Phe) in body fluids may be assayed using the compsns.
 XX
 SQ Sequence 10475 BP; 2415 A; 3060 C; 2942 G; 2056 T; 2 other;

Alignment Scores:
 Pred. No.: 215 Length: 10475
 Score: 52.00 Matches: 10
 Percent Similarity: 76.47% Conservative: 3
 Best Local Similarity: 58.82% Mismatches: 4
 Query Match: 50.00% Indels: 0
 DB: 9 Gaps: 0

US-09-833-017B-4 (1-21) x AAN80643 (1-10475)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGln 17
 DB 543 TCTGGTCTATGCACATGGTTGCTTAGTCAGAGAGCATTTACAA 493

RESULT 9
 ABL67064/C
 ID ABL67064 standard; DNA; 10476 BP.
 XX ABL67064;
 XX
 DT 15-MAY-2002 (first entry)
 DE Thyroid cancer related gene sequence SEQ ID NO:5401.
 XX
 KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
 KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200194629-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 30-MAY-2001; 2001WO-US10838.
 XX
 PR 05-JUN-2000; 2000US-209473P.
 PR 05-JUN-2000; 2000US-209531P.
 PR 18-SEP-2000; 2000US-233133P.
 PR 18-SEP-2000; 2000US-233617P.
 PR 20-SEP-2000; 2000US-234009P.
 PR 20-SEP-2000; 2000US-234034P.
 PR 22-SEP-2000; 2000US-234052P.
 PR 22-SEP-2000; 2000US-234509P.
 PR 22-SEP-2000; 2000US-234567P.
 PR 25-SEP-2000; 2000US-234923P.
 PR 25-SEP-2000; 2000US-234924P.
 PR 25-SEP-2000; 2000US-235077P.
 PR 25-SEP-2000; 2000US-235082P.

PR 25-SEP-2000; 2000US-235134P.
 PR 25-SEP-2000; 2000US-235280P.
 PR 26-SEP-2000; 2000US-235637P.
 PR 26-SEP-2000; 2000US-235638P.
 PR 27-SEP-2000; 2000US-235711P.
 PR 27-SEP-2000; 2000US-235720P.
 PR 27-SEP-2000; 2000US-235840P.
 PR 27-SEP-2000; 2000US-235863P.
 PR 28-SEP-2000; 2000US-236028P.
 PR 28-SEP-2000; 2000US-236032P.
 PR 28-SEP-2000; 2000US-236033P.
 PR 28-SEP-2000; 2000US-236034P.
 PR 28-SEP-2000; 2000US-236109P.
 PR 29-SEP-2000; 2000US-236111P.
 PR 29-SEP-2000; 2000US-236842P.
 PR 29-SEP-2000; 2000US-236891P.
 PR 02-OCT-2000; 2000US-237172P.
 PR 02-OCT-2000; 2000US-237173P.
 PR 02-OCT-2000; 2000US-237278P.
 PR 02-OCT-2000; 2000US-237294P.
 PR 02-OCT-2000; 2000US-237295P.
 PR 02-OCT-2000; 2000US-237316P.
 PR 03-OCT-2000; 2000US-237425P.
 PR 03-OCT-2000; 2000US-237598P.
 PR 03-OCT-2000; 2000US-237604P.
 PR 03-OCT-2000; 2000US-237606P.
 PR 03-OCT-2000; 2000US-237608P.
 PR 01-NOV-2000; 2000US-244867P.
 PR 01-NOV-2000; 2000US-245084P.
 XX
 PA (AVAL-) AVALON PHARM.
 XX
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
 PI Soppet DR, Weaver Z;
 XX WPI; 2002-188264/24.
 XX

Screening for anti-neoplastic agent involves exposing cells to a
 chemical agent to be tested for anti-neoplastic activity, and
 determining a change in expression of a gene of a signature gene set -
 Claim 1; SEQ ID 5401; 44pp; English.

The present invention describes a method (M1) for screening for an
 anti-neoplastic agent. The method involves exposing cells to a chemical
 agent to be tested for anti-neoplastic activity, determining a change in
 expression of at least one gene (I) of a signature gene set, where (I)
 comprises a sequence (S) selected from 8447 sequences (given in ABL61664
 to ABL70110), or is at least 95% identical to (S), where a change in
 expression is indicative of anti-neoplastic activity. (I) has cytostatic
 activity and can be used in gene therapy. M1 can be used for screening
 an anti-neoplastic agent, and can be used for producing a product which
 is the data collected with respect to the anti-neoplastic agent as a
 result of M1, and the data is sufficient to convey the chemical
 structure and/or properties of the agent. M1 can be used in the
 treatment of cancer such as colon, breast, stomach, lung, thyroid,
 oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 carcinoma, papillary carcinoma and Wilm's tumour.

SQ Sequence 10476 BP; 2416 A; 3059 C; 2944 G; 2057 T; 0 other;

Alignment Scores:
 Pred. No.: 215 Length: 10476
 Score: 52.00 Matches: 10
 Percent Similarity: 76.47% Conservative: 3
 Best Local Similarity: 58.82% Mismatches: 4
 Query Match: 50.00% Indels: 0
 DB: 24 Gaps: 0

US-09-833-017B-4 (1-21) x ABL67064 (1-10476)

XX DR WPI; 2001-616771/71.
 XX PT New polynucleotide for treating Chlamydia infections encodes a
 PT polynucleotides containing an immunogenic portion of a Chlamydia
 PT antigen -
 XX
 XX
 XX Claim 1; Page 116-117; 208pp; English.
 XX
 XX The invention relates to isolated polynucleotide encoding at least
 CC a partial Chlamydia protein which is an antigenic fragment, or the
 CC complements, fragments, homologues and variants, and antibodies
 CC raised against the antigenic proteins (or fragments). The nucleic
 CC acids, proteins and antibodies are used to diagnose and treat Chlamydia
 CC infections (e.g. a sexually transmitted disease, pelvic inflammatory
 CC disease (PID), acute respiratory tract infection, trachoma,
 CC atherosclerosis and coronary heart disease) in a patient, and in
 CC the treatment of male infertility. The compounds of the invention are
 CC also useful for detecting the presence of Chlamydia in a patient, and
 CC stimulating and/or expanding T cells specific for a Chlamydia protein.
 CC The present sequence encodes a Chlamydia antigen.
 XX
 XX Sequence 2017 BP; 490 A; 443 C; 389 G; 695 T; 0 other;

Alignment Scores:
 Pred. No.: 64.3 Length: 2017
 Score: 50.00 Matches: 10
 Percent Similarity: 78.57% Conservative: 1
 Best Local Similarity: 71.43% Mismatches: 3
 Query Match: 48.08% Indels: 0
 DB: 22 Gaps: 0

US-09-833-017B-4 (1-21) x AAS56980 (1-2017)

Qy 3 SerLeuSerThrPhePheArgLeuPheAsnArgSerPheThr 16
 Db 353 GCACCTAATGACCTTTTCAGTTTATTTAATAGTCTCTTAACA 394

RESULT 12
 ABL04050
 ID ABL04050 standard; cDNA; 2932 BP.
 XX
 XX ABL04050;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 6632.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW
 XX Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 XX
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB59947.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

PS Claim 1; SEQ ID NO 6632; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 2932 BP; 859 A; 603 C; 591 G; 879 T; 0 other;

Alignment Scores:
 Pred. No.: 102 Length: 2932
 Score: 50.00 Matches: 10
 Percent Similarity: 80.00% Conservative: 2
 Best Local Similarity: 66.67% Mismatches: 3
 Query Match: 48.08% Indels: 0
 DB: 23 Gaps: 0

US-09-833-017B-4 (1-21) x ABL04050 (1-2932)

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 15
 Db 2127 TCAGGATCGTTCAGTTTATTTCTATTTCTCGGGATTTT 2171

RESULT 13
 ABL03962
 ID ABL03962 standard; cDNA; 3596 BP.
 XX
 XX ABL03962;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 6368.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 KW
 XX Drosophila melanogaster.
 XX
 XX WO200171042-A2.
 XX
 XX 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX 23-MAR-2000; 2000US-191637P.
 XX
 XX 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR P-PSDB; ABB59859.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX

Claim 1; SEQ ID NO 6368; 21pp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC antibodies in standard immunoassays, especially for diagnosing or
CC monitoring infections. Antibodies which bind the protein are used to
CC detect corresponding antigens, to purify the protein and for passive
CC immunisation (optionally coupled to a toxin). Vaccines are administered,
CC e.g. by injection, orally or through the skin, typically at 0.01-1000
CC (especially 10-300) mu g/ml per dose.

XX
SQ Sequence 694 BP; 271 A; 77 C; 111 G; 234 T; 1 other;

Alignment Scores: 26.1 Length: 694
Pred. No.: 49.00 Matches: 10
Score: 70.59% Conservatives: 2
Percent Similarity: 58.82% Mismatches: 5
Best Local Similarity: 47.12% Indels: 0
Query Match: 19 Gaps: 0
DB:

US-09-833-017B-4 (1-21) x AAV27411 (1-694)

OY 5 SerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
 ||||| ||| ::||| ::||| |||||
DB 210 TCTACCTTTCTCGGTTTTCACATAAAAGTTTCACGGCTTCTCTAGGAAAA 160

RESULT 15
ABQ84879/c
ID ABQ84879 standard; DNA; 694 BP.
XX
AC ABQ84879;
XX
DT 04-SEP-2002 (first entry)
XX
DE S. pneumoniae SP100 nucleotide sequence SEQ ID NO:175.
XX
KW Streptococcus pneumoniae; epitope; vaccine; antigenic protein;
KW antibacterial; Streptococcal infection; detection; gene; ds.
XX
OS Streptococcus pneumoniae.
XX
PN US2002061545-A1.
XX
PD 23-MAY-2002.
XX
PF 22-JAN-2001; 2001US-0765272.
XX
PR 30-OCT-1997; 97US-0961083.
XX
PA (CHOI/) CHOI G H.
PA (KUNS/) KUNSCH C A.
PA (BARA/) BARASH S C.
PA (DILL/) DILLON P J.
PA (DOUG/) DOUGHERTY B.
PA (FANN/) FANNON M R.
PA (ROSE/) ROSEN C A.
XX
PI Choi GH, Kunsch CA, Barash SC, Dillon PJ, Dougherty B, Fannon MR;
PI Rosen CA;
XX
DR WPI; 2002-479261/51.
DR P-PSDB; ABP54644.
XX
XX New Streptococcus pneumoniae antigens, useful for detecting
PT Streptococcus and for preventing or attenuating disease caused by
PT Streptococcus infection -
XX
PS Claim 1; Page 44; 70pp; English.

CC ABQ84792 to ABQ84904 represents nucleic acids which encode the
CC Streptococcus pneumoniae antigens given in ABP54557 to ABP54669.
CC The S. pneumoniae antigens have antibacterial activity and can be
CC used in vaccines. The S. pneumoniae antigens can also be used to
CC prevent or attenuate a Streptococcal infection in an animal. The
CC polynucleotides encoding the S. pneumoniae antigens can be used to
CC detect Streptococcus nucleic acids. ABQ84905 to ABQ85130 represent

CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 3596 BP; 898 A; 897 C; 809 G; 992 T; 0 other;

Alignment Scores: 131 Length: 3596
Pred. No.: 50.00 Matches: 10
Score: 80.00% Conservatives: 2
Percent Similarity: 66.67% Mismatches: 3
Best Local Similarity: 48.08% Indels: 0
Query Match: 23 Gaps: 0
DB:

US-09-833-017B-4 (1-21) x ABL03962 (1-3596)

OY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 15
 ||||| ||||| ||||| |||||
DB 831 TCAGGATCGTGGTATATATTCTATTTATTTCTCGGGATTTT 875

RESULT 14
AAV27411/c
ID AAV27411 standard; DNA; 694 BP.
XX
AC AAV27411;
XX
DT 02-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae SP0100 nucleotide.
XX
KW Streptococcus pneumoniae; antigen; vaccine; infection; diagnosis;
KW detection; pneumonia; otitis media; meningitis; ss.
XX
OS Streptococcus pneumoniae.
XX
FH Key Location/Qualifiers
FT CDS 2..694
FT /tag= a
FT /product= "SP0100"
FT /transl_except= (pos:590..592,aa:Xaa)
FT /note= "no stop codon given; Xaa is unspecified"

XX
PN WO9818930-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US19422.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Choi GH, Hromockyj A, Johnson LS, Kunsch CA;
PI WPI; 1998-272224/24.
DR P-PSDB; AAW61225.
XX
XX Nucleic acid encoding antigenic peptide(s) from Streptococcus
PT pneumoniae - or their epitope-containing fragments, useful in
PT protective or therapeutic vaccines, and for diagnosis
XX
PS Claim 1; Page 84; 118pp; English.

CC The present sequence encodes a protein from Streptococcus pneumoniae.
CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
CC can be useful in vaccines for inducing protective antibodies against
CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
CC are used to detect Streptococcus infection (by usual hybridisation or
CC amplification methods), also for isolating Streptococcus genes or their
CC allelic variants. The protein can be used similarly to detect specific

primers used in the cloning of *S. pneumoniae* ORFs (open reading frames) which are used in an example from the present invention.

CC
XX
WHICH ARE USED IN AN EXAMPLE FROM THE PRESENT INVENTION.

XX

Sequence 694 BP; 271 A; 77 C; 111 G; 234 T; 1 other;

Alignment Scores:

Fragmented Scores:					
Pred. No.:	26.1	Length:	694		
Score:	49.00	Matches:	10		
Percent Similarity:	70.59%	Conservative:	2		
Best Local Similarity:	58.82%	Mismatches:	5		
Query Match:	47.11%	Indels:	0		
DB:	24	Gaps:	0		

tried: NO.:
 Score:

Percent Similarity:

Best Local Similarity

Query Match:

Query Malcol.
DB: 24

Gaps:

US-09-833-017B-4 (1-21) x ABQ84879 (1-694)

Qy 5 SerThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21

Db 210 TCTACCTTTTCTCCGTTTTTCTCACTAAAAGTTTCACGGCTTCTCTAGGAAA

Search completed: November 8, 2002, 22:11:00
Job time : 108.687 secs

Job time : 108.687 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: November 8, 2002, 22:05:37 ; Search time 21.9403 Seconds
(without alignments)
293.533 Million cell updates/sec

Title: US-09-833-017B-4
Perfect score: 104
Sequence: 1 SCSLSFFRLNRSTQALGK 21

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q/cgn2_1/USPTO.spool/US09833017/runat_05112002_105349_4858/app_query.fasta_1.398
-DB-Issued_Patents_NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017.cgn2_1_17_runat_05112002_105349_4858 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
c 1	49	47.1	694	3	US-08-961-083-175	Sequence 175, Appl
c 2	49	47.1	2756	1	US-08-187-793-1	Sequence 1, Appl
c 3	49	47.1	3318	1	US-08-187-793-1	Sequence 3, Appl
c 4	49	47.1	6004	4	US-08-961-527-27	Sequence 27, Appl
c 5	48	46.5	1317	1	US-09-134-001C-302	Sequence 302, Appl
c 6	46.5	44.7	2682	1	US-07-971-819A-1	Sequence 1, Appl
c 7	46.5	44.7	2682	1	US-07-977-434-3	Sequence 3, Appl
c 8	46.5	44.7	2682	1	US-08-475-231-1	Sequence 1, Appl
c 9	46.5	44.7	2682	1	US-08-458-819-3	Sequence 3, Appl
c 10	46.5	44.7	2682	1	US-09-105-697-9	Sequence 9, Appl
c 11	46.5	44.7	2682	5	PCT-US91-07035-3	Sequence 3, Appl
c 12	45	43.3	1651	4	US-09-484-970B-17	Sequence 17, Appl

c 13	45	43.3	8411	4	US-08-961-527-16	Sequence 16, Appl
c 14	45	43.3	65042	4	US-09-784-316-3	Sequence 3, Appl
c 15	44	42.3	1179	4	US-08-745-995A-22	Sequence 22, Appl
c 16	44	42.3	1179	4	US-08-745-995A-23	Sequence 23, Appl
c 17	44	42.3	1191	4	US-08-745-995A-4	Sequence 4, Appl
c 18	44	42.3	1191	4	US-08-745-995A-5	Sequence 5, Appl
c 19	44	42.3	1191	4	US-08-745-995A-34	Sequence 34, Appl
c 20	44	42.3	1191	4	US-08-745-995A-35	Sequence 35, Appl
c 21	44	42.3	1194	4	US-08-745-995A-28	Sequence 28, Appl
c 22	44	42.3	1194	4	US-08-745-995A-29	Sequence 29, Appl
c 23	44	42.3	1197	4	US-08-745-995A-10	Sequence 10, Appl
c 24	44	42.3	1197	4	US-08-745-995A-11	Sequence 11, Appl
c 25	44	42.3	1260	4	US-08-745-995A-16	Sequence 16, Appl
c 26	44	42.3	1260	4	US-08-745-995A-17	Sequence 17, Appl
c 27	44	42.3	1358	4	US-08-745-995A-7	Sequence 7, Appl
c 28	44	42.3	1358	4	US-08-745-995A-9	Sequence 9, Appl
c 29	44	42.3	1414	4	US-08-745-995A-19	Sequence 19, Appl
c 30	44	42.3	1414	4	US-08-745-995A-21	Sequence 21, Appl
c 31	44	42.3	1454	4	US-08-745-995A-31	Sequence 31, Appl
c 32	44	42.3	1454	4	US-08-745-995A-33	Sequence 33, Appl
c 33	44	42.3	1492	4	US-08-745-995A-25	Sequence 25, Appl
c 34	44	42.3	1492	4	US-08-745-995A-27	Sequence 27, Appl
c 35	44	42.3	1584	4	US-08-745-995A-1	Sequence 1, Appl
c 36	44	42.3	1584	4	US-08-745-995A-3	Sequence 3, Appl
c 37	44	42.3	1838	4	US-08-745-995A-13	Sequence 13, Appl
c 38	44	42.3	1838	4	US-08-745-995A-15	Sequence 15, Appl
c 39	43	41.3	510	4	US-09-134-001C-1753	Sequence 1753, Ap
c 40	43	41.3	777	4	US-08-809-326A-7	Sequence 7, Appl
c 41	43	41.3	813	4	US-08-809-326A-4	Sequence 4, Appl
c 42	43	41.3	1000	4	US-09-641-638-647	Sequence 647, App
c 43	43	41.3	1048	4	US-08-809-326A-9	Sequence 9, Appl
c 44	43	41.3	1296	4	US-08-809-326A-18	Sequence 18, Appl
c 45	43	41.3	1464	4	US-08-809-326A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-961-083-175/c
; Sequence 175, Application US/08961083
; Patent No. 6159469
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,083
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 175:
; SEQUENCE CHARACTERISTICS:

Best Local Similarity: 83.33% Mismatches: 1
Query Match: 47.12% Indels: 0
DB: 1 Gaps: 0

US-09-833-017B-4 (1-21) x US-08-187-793-3 (1-3318)

QY 4 LeuSerThrPheArgLeuPheAsnArgSerPhe 15
|||||
Db 2758 TTATCTACCTTTGTCAGGCTTTTGACAGGTCATT 2793

RESULT 4

US-09-961-527-27
; Sequence 27, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6004 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-27

Alignment Scores:
Pred. No.: 76.2 Length: 6004
Score: 49.00 Matches: 10
Percent Similarity: 70.59% Conservative: 2
Best Local Similarity: 58.82% Mismatches: 5
Query Match: 47.12% Indels: 0
DB: 4 Gaps: 0

US-09-833-017B-4 (1-21) x US-08-961-527-27 (1-6004)

QY 5 SerThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
|||||
Db 418 TCTACCTTTCTCCGTTTTCATAAAAGTTTCACGCGTCTCTAGGAAA 468

RESULT 5

US-09-134-001C-302/c
; Sequence 302, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 302
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-302

Alignment Scores:
Pred. No.: 17.9 Length: 1317
Score: 48.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 46.15% Indels: 0
DB: 4 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-134-001C-302 (1-1317)

QY 7 PhePheArgLeuPheAsnArgSerPhe 15

|||||
Db 54 TTTTTCGCTGTGTTTATCGTAGTTT 28

RESULT 6

US-07-971-819A-1/c
; Sequence 1, Application US/07971819A
; Patent No. 5420029
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David H.
; APPLICANT: Lawyer, Frances C.
; APPLICANT: Stoffel, Susanne
; TITLE OF INVENTION: Purified Thermostable Nucleic Acid
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/971,819A
; FILING DATE: 19930203
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stacey R. Sias, Ph.D.
; REGISTRATION NUMBER: 32,630
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2863
; TELEFAX: (510) 814-2977
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS

LOCATION: 1..2682
US-07-971-819A-1
Alignment Scores: 78.3 Length: 2682
Pred. No.: 46.50 Matches: 12
Score: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 8
Query Match: 44.71% Indels: 1
DB: 1 Gaps: 1
US-09-833-017B-4 (1-21) x US-07-971-819A-1 (1-2682)
QY 1 SerGlySerLeuSerThrPhePheArgLeuPhe---AsnArgSerPheThrGlnAlaLeu 19
Db 1134 TCCGGGGTCTCCAGAAATTTCTTTGAGCTTTTTCAGAACCTCTTTTTCGCCAGGTTCTG 1075
QY 20 Gly 20
Db 1074 GGC 1072
RESULT 7
US-07-977-434-3/C
Sequence 3, Application US/07977434
Patent No. 5466591
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Abramson, Richard D.
TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
THERMOSTABLE DNA POLYMERASES
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: 7
SOFTWARE: WordPerfect 2.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/977,434
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,490
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,466
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 590,213
FILING DATE: 28-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 523,394
FILING DATE: 15-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 143,441
FILING DATE: 12-JAN-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 063,509
FILING DATE: 17-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 899,241
FILING DATE: 22-AUG-1986
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 746,121
FILING DATE: 15-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US90/07641
FILING DATE: 21-DEC-1990

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 585,471
FILING DATE: 20-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 455,611
FILING DATE: 22-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 609,157
FILING DATE: 02-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 557,517
FILING DATE: 24-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Luann Cserr
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: Case No. 5466591 8753
TELEPHONE: (510) 814-2972
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2682 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Thermotoga maritima
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2679
US-07-977-434-3
Alignment Scores: 78.3 Length: 2682
Pred. No.: 46.50 Matches: 12
Score: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 8
Query Match: 44.71% Indels: 1
DB: 1 Gaps: 1
US-09-833-017B-4 (1-21) x US-07-977-434-3 (1-2682)
QY 1 SerGlySerLeuSerThrPhePheArgLeuPhe---AsnArgSerPheThrGlnAlaLeu 19
Db 1134 TCCGGGGTCTCCAGAAATTTCTTTGAGCTTTTTCAGAACCTCTTTTTCGCCAGGTTCTG 1075
QY 20 Gly 20
Db 1074 GGC 1072
RESULT 8
US-08-475-231-1/C
Sequence 1, Application US/08475231
Patent No. 5624833
GENERAL INFORMATION:
APPLICANT: Gelfand, David H.
APPLICANT: Lawyer, Frances C.
APPLICANT: Stoffel, Susanne
TITLE OF INVENTION: Purified Thermostable Nucleic Acid
TITLE OF INVENTION: Polymerase Enzyme from Thermotoga Maritima
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: USA
ZIP: 07110-1199
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb storage
COMPUTER: Apple Macintosh

```

; OPERATING SYSTEM: Macintosh 7
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,231
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/971,819
; FILING DATE: 03-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stacey R. Sias, Ph.D.
; REGISTRATION NUMBER: 32,630
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2863
; TELEFAX: (510) 814-2977
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2682
; US-08-475-231-1

```

```

Alignment Scores:
Pred. No.: 78.3 Length: 2682
Score: 46.50 Matches: 12
Percent Similarity: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 8
Query Match: 44.71% Indels: 1
DB: 1 Gaps: 1

```

```

US-09-833-017B-4 (1-21) x US-08-475-231-1 (1-2682)

```

```

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPhe---AsnArgSerPheThrGlnAlaLeu 19
Db 1134 TCCGGGTCCTCCAGATTCTTTGAGCTTTTTCAGAACCTCTTTTTCGTCGAGTTCTG 1075

```

```

Qy 20 Gly 20

```

```

Db 1074 GGC 1072

```

```

RESULT 9

```

```

US-08-458-819-3/c

```

```

; Sequence 3, Application US/08458819
; Patent No. 5795762
; GENERAL INFORMATION:
; APPLICANT: Gelfand, David H.
; APPLICANT: Abramson, Richard D.
; TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF
; TITLE OF INVENTION: THERMOSTABLE DNA POLYMERASES
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; ZIP: 07110-1199
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: 7
; SOFTWARE: WordPerfect 2.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,819
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: 07/977,434
; FILING DATE: 23-FEB-1993
; APPLICATION NUMBER: US 590,490
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,466
; FILING DATE: 28-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 590,213
; FILING DATE: 28-SEP-1990
; APPLICATION NUMBER: US 523,394
; FILING DATE: 15-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 143,441
; FILING DATE: 12-JAN-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 063,509
; FILING DATE: 17-JUN-1987
; APPLICATION NUMBER: US 899,241
; FILING DATE: 22-AUG-1986
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 746,121
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US90/07641
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 585,471
; FILING DATE: 20-SEP-1990
; APPLICATION NUMBER: US 455,611
; FILING DATE: 22-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 609,157
; FILING DATE: 02-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 557,517
; FILING DATE: 24-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Luann Cserr
; REGISTRATION NUMBER: 31,822
; REFERENCE/DOCKET NUMBER: Case No. 5795762 8753
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 814-2972
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Thermotoga maritima
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2679
; US-08-458-819-3

```

```

Alignment Scores:
Pred. No.: 78.3 Length: 2682
Score: 46.50 Matches: 12
Percent Similarity: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 8
Query Match: 44.71% Indels: 1
DB: 1 Gaps: 1

```

```

US-09-833-017B-4 (1-21) x US-08-458-819-3 (1-2682)

```

```

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPhe---AsnArgSerPheThrGlnAlaLeu 19

```

TITLE OF INVENTION: 5' TO 3' EXONUCLEASE MUTATIONS OF

```
;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2679
PCT-US91-07035-3
Alignment Scores:
Pred. No.: 78.3 Length: 2682
Score: 46.50 Matches: 12
Percent Similarity: 57.14% Conservative: 0
Best Local Similarity: 57.14% Mismatches: 8
Query Match: 44.71% Indels: 1
DB: 5 Gaps: 1

US-09-833-017b-4 (1-21) x PCT-US91-07035-3 (1-2682)
Qy 1 SerGlySerLeuSerThrPheArgLeuPhe--AsnArgSerPheThrGlnAlaLeu 19
Db 1134 TCCGGGTCCTCCAGAAATTCITTAGCTTTTCAGAACCTCTTTTCCTCCAGGTTCTG 1075
Qy 20 Gly 20
Db 1074 GGC 1072

RESULT 12
US-09-484-970B-17/c
; Sequence 17, Application US/09484970B
; Patent No. 6426186
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen A.
; APPLICANT: Volkmut, Wayne
; APPLICANT: Walker, Michael G.
; TITLE OF INVENTION: BONE REMODELING GENES
; FILE REFERENCE: PB-0014 US
; CURRENT APPLICATION NUMBER: US/09/484,970B
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PERL Program
; SEQ ID NO 17
; LENGTH: 1651
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6426186 126510.2CBI
; NAME/KEY: unsure
; LOCATION: 767-846
; OTHER INFORMATION: a, t, c, g, or other
US-09-484-970B-17
Alignment Scores:
Pred. No.: 79.4 Length: 1651
Score: 45.00 Matches: 9
Percent Similarity: 76.92% Conservative: 1
Best Local Similarity: 69.23% Mismatches: 3
Query Match: 43.27% Indels: 0
DB: 4 Gaps: 0

US-09-833-017b-4 (1-21) x US-09-484-970B-17 (1-1651)
Qy 4 LeuSerThrPheArgLeuPheAsnArgSerPheThr 16
Db 519 GTGAGCACCTCTTCCGGTCTTCCCGCGCTCTCTGACC 481

RESULT 13
US-08-961-527-16/c
; Sequence 16, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340Pl
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8411 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-961-527-16
Alignment Scores:
Pred. No.: 581 Length: 8411
Score: 45.00 Matches: 8
Percent Similarity: 81.25% Conservative: 5
Best Local Similarity: 50.00% Mismatches: 3
Query Match: 43.27% Indels: 0
DB: 4 Gaps: 0

US-09-833-017b-4 (1-21) x US-08-961-527-16 (1-8411)
Qy 6 ThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
Db 3443 ACCCTCTATCGTTAGATACCAAACTTTCACCGAGGCAATCAAAA 3396

RESULT 14
US-09-784-316-3/c
; Sequence 3, Application US/09784316
; Patent No. 6461843
; GENERAL INFORMATION:
; APPLICANT: Wei, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01139
; CURRENT APPLICATION NUMBER: US/09/784,316
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 65042
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(65042)
; OTHER INFORMATION: n = A,T,C or G
US-09-784-316-3
Alignment Scores:
Pred. No.: 7.07e+03 Length: 65042
Score: 45.00 Matches: 8
```

wed Nov 13 10:42:17 2002

Percent Similarity: 72.22% Conservative: 5
Best Local Similarity: 44.44% Mismatches: 5
Query Match: 43.27% Indels: 0
DB: 4 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-784-316-3 (1-65042)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAla 18
||||| ||||| ::|||:: ||| :::::|||||:::
DB 42098 TCAGGAATCTCAAAATGTTACAGAATAATTATACCAATTATACCAATCT 42045

RESULT 15

US-08-745-995A-22/c
; Sequence 22, Application US/08745995A
; Patent No. 6372887
; GENERAL INFORMATION:
; APPLICANT: Silver, Gary M.
; APPLICANT: Wisniewski, Nancy
; TITLE OF INVENTION: No. 6372887el Serine Protease Inhibitor
; TITLE OF INVENTION: Nucleic Acid Molecules, Proteins
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA: US/08/745,995A
; APPLICATION NUMBER: US/08/745,995A
; FILING DATE: 07-NOV-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: FC-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1179 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-745-995A-22

Alignment Scores:
Pred. No.: 78.9 Length: 1179
Score: 44.00 Matches: 10
Percent Similarity: 66.67% Conservative: 0
Best Local Similarity: 66.67% Mismatches: 5
Query Match: 42.31% Indels: 0
DB: 4 Gaps: 0

US-09-833-017B-4 (1-21) x US-08-745-995A-22 (1-1179)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPhe 15
||| ||| ||||| ||||| ||||| ||||| |||||
DB 1041 AGCTGACGCTTCAGCACCTCTCTTACTTCAATGAAGCTTT 997

Search completed: November 8, 2002, 23:41:22
Job time : 28.9403 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 22:57:42 : Search time 20.6866 Seconds
(without alignments)
360.161 Million cell updates/sec

Title: US-09-833-017B-4

Perfect score: 104

Sequence: 1 GSSLTFRLEFRSFTQALGK 21

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 320260 seqs, 177392727 residues

Total number of hits satisfying chosen parameters: 640520

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09833017/runat_05112002_105351_4940/app_query.fasta_1.398
-DB=Published Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cgi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09833017.ecgn_1_18_srunat_05112002_105351_4940
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	141	10	US-09-833-017-1
2	104	100.0	2557	10	US-09-833-017-21
c 3	52	50.0	10476	10	US-09-964-824A-98
c 4	52	50.0	10476	10	US-09-964-824A-552

50	48.1	2017	12	US-10-007-693-12	Sequence 12, Appl
49	47.1	694	10	US-09-765-272-175	Sequence 175, App
47	45.2	870	10	US-09-770-445-564	Sequence 564, App
46	44.2	48667	10	US-09-822-268A-3	Sequence 3, Appli
c 8	45	284	10	US-09-864-761-32659	Sequence 32659, A
45	43.3	333	10	US-09-917-800A-276	Sequence 276, App
45	43.3	430	10	US-09-728-446-245	Sequence 245, App
45	43.3	489	10	US-09-733-387-11	Sequence 11, Appl
c 12	45	525	10	US-09-974-300-6233	Sequence 6233, Ap
c 13	45	528	10	US-09-733-387-31	Sequence 31, Appl
c 14	45	584	10	US-09-864-761-16165	Sequence 16165, A
c 15	45	612	10	US-09-733-387-51	Sequence 51, Appl
c 16	45	702	10	US-09-733-387-9	Sequence 9, Appli
c 17	45	741	10	US-09-733-387-29	Sequence 29, Appl
c 18	45	825	10	US-09-733-387-49	Sequence 49, Appl
c 19	45	897	10	US-09-733-387-5	Sequence 5, Appli
c 20	45	936	10	US-09-733-387-25	Sequence 25, Appl
c 21	45	942	10	US-09-974-300-2095	Sequence 2095, Ap
c 22	45	1020	10	US-09-733-387-45	Sequence 45, Appl
c 23	45	1080	10	US-09-733-387-7	Sequence 7, Appli
c 24	45	1119	10	US-09-733-387-27	Sequence 27, Appl
c 25	45	1203	10	US-09-733-387-47	Sequence 47, Appl
c 26	45	1527	10	US-09-733-387-3	Sequence 3, Appli
c 27	45	1539	10	US-09-887-377-1	Sequence 1, Appli
c 28	45	1566	10	US-09-733-387-23	Sequence 23, Appl
c 29	45	1650	10	US-09-733-387-43	Sequence 43, Appl
c 30	45	4036	10	US-09-733-387-53	Sequence 53, Appl
c 31	45	4036	10	US-09-954-456-1996	Sequence 1996, Ap
c 32	45	348	10	US-09-770-791-742	Sequence 742, App
c 33	44	599	10	US-09-864-761-12681	Sequence 12681, A
c 34	44	1053	10	US-09-833-381-365	Sequence 365, App
c 35	44	1405	10	US-09-943-272-1	Sequence 1, Appli
c 36	44	32248	10	US-09-764-864-1769	Sequence 1769, Ap
c 37	44	32248	10	US-09-764-877-3487	Sequence 3487, Ap
c 38	44	32248	10	US-09-813-320-3	Sequence 3, Appli
c 39	44	397658	10	US-09-813-320-3	Sequence 3, Appli
c 40	44	640681	10	US-09-790-988-1	Sequence 1, Appli
c 41	43.5	438	10	US-09-960-352-9083	Sequence 9083, Ap
c 42	43	215	10	US-09-815-242-2813	Sequence 2813, Ap
c 43	43	344	10	US-09-815-242-2237	Sequence 2237, Ap
c 44	43	439	10	US-09-563-817-724	Sequence 724, App
c 45	43	501	10	US-09-815-242-4689	Sequence 4689, Ap

ALIGNMENTS

RESULT 1
US-09-833-017-1
; Sequence 1, Application US/09833017
; Patent No. US20020081302A1
; GENERAL INFORMATION:
; APPLICANT: CVITKOVITCH, Dennis
; TITLE OF INVENTION: SIGNAL PEPTIDES NUCLEIC ACID MOLECULES AND METHODS FOR TREATM
; FILE REFERENCE: P04885U1
; CURRENT APPLICATION NUMBER: US/09/833,017
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: PCT/CA00/00605
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 2,302,861
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 2,332,733
; PRIOR FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 141
; TYPE: DNA
; ORGANISM: Streptococcus mutans
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(141)
US-09-833-017-1

; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
 ; TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
 ; FILE REFERENCE: 210121.515C2
 ; CURRENT APPLICATION NUMBER: US/10/007,693
 ; CURRENT FILING DATE: 2001-12-05
 ; NUMBER OF SEQ ID NOS: 157
 ; SEQ ID NO 12
 ; LENGTH: 2017
 ; TYPE: DNA
 ; ORGANISM: Chlamydia trachomatis
 US-10-007-693-12

Alignment Scores:
 Pred. No.: 11.9 Length: 2017
 Score: 50.00 Matches: 10
 Percent Similarity: 78.5% Conservative: 1
 Best Local Similarity: 71.43% Mismatches: 3
 Query Match: 48.08% Indels: 0
 DB: 12 Gaps: 0

US-09-833-017b-4 (1-21) x US-10-007-693-12 (1-2017)

QY 3 SerLeuSerThrpPhePheArgLeuPheAsnArgSerPheThr 16
 Db 353 GCCTAATGACCTTTTTCAGTTTATTAATAGTCTTTAACA 394

RESULT 6

US-09-765-272-175/c

Sequence 175, Application US/09765272

Patent No. US20020061545A1

GENERAL INFORMATION:

APPLICANT: Choi et. al.

TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines

NUMBER OF SEQUENCES: 452

CORRESPONDENCE ADDRESS:

ADDRESSER: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/765,272

FILING DATE: 22-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/961,083

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 175:

SEQUENCE CHARACTERISTICS:

LENGTH: 694 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 175:

US-09-765-272-175

Alignment Scores:

Pred. No.: 4.82 Length: 694
 Score: 49.00 Matches: 10
 Percent Similarity: 70.59% Conservative: 2
 Best Local Similarity: 58.82% Mismatches: 5
 Query Match: 47.12% Indels: 0
 DB: 10 Gaps: 0

US-09-833-017b-4 (1-21) x US-09-765-272-175 (1-694)

QY 5 SerThrpPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
 Db 210 TCTACCTTTTCTCCGTTTTCACATAAAAGTTTCACGGCTTCTTAGGAAA 160

RESULT 7

US-09-770-445-564

Sequence 564, Application US/09770445

Patent No. US20020023281A1

GENERAL INFORMATION:

APPLICANT: Gorlach, Jorn

APPLICANT: An, Yong-Qiang

APPLICANT: Hamilton, Carol M.

APPLICANT: Price, Jennifer L.

APPLICANT: Raines, Tracy M.

APPLICANT: Yu, Yang

APPLICANT: Rameaka, Joshua G.

APPLICANT: Page, Amy

APPLICANT: Matthew, Abraham V.

APPLICANT: Ledford, Brooke L.

APPLICANT: Woessner, Jeffrey P.

APPLICANT: Haas, William David

APPLICANT: Garcia, Carlos A.

APPLICANT: Kriker, Maja

APPLICANT: Slader, Ted

APPLICANT: Davis, Keith R.

APPLICANT: Allen, Keith

APPLICANT: Hoffman, Neil

APPLICANT: Hurban, Patrick

TITLE OF INVENTION: Expressed Sequences of Arabidopsis

FILE REFERENCE: 2023US (PARA-012PRV)

CURRENT APPLICATION NUMBER: US/09/770,445

CURRENT FILING DATE: 2001-01-26

PRIOR APPLICATION NUMBER: US 60/178,472

PRIOR FILING DATE: 2000-01-27

NUMBER OF SEQ ID NOS: 999

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 564

LENGTH: 870

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-770-445-564

Alignment Scores:

Pred. No.: 14.2 Length: 870
 Score: 47.00 Matches: 9
 Percent Similarity: 65.00% Conservative: 4
 Best Local Similarity: 45.00% Mismatches: 7
 Query Match: 45.19% Indels: 0
 DB: 10 Gaps: 0

US-09-833-017b-4 (1-21) x US-09-770-445-564 (1-870)

QY 2 GlySerLeuSerThrpPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
 Db 785 GGTATTATCAAACTCTTTTGTCTTTTCGCCAGACATTTACAGAACCTTTTCGAAAA 844

RESULT 8

US-09-822-268A-3/c

Sequence 3, Application US/09822268A

Patent No. US20020048787A1

GENERAL INFORMATION:

APPLICANT: Wei, Ming-Hui et al.

TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

```
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/822,268A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/211,387
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 48667
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(48667)
; OTHER INFORMATION: n = A,T,C or G
US-09-822-268A-3

Alignment Scores:
Pred. No.: 2.89e+03 Length: 48667
Score: 46.00 Matches: 9
Percent Similarity: 72.22% Conservative: 4
Best Local Similarity: 50.00% Mismatches: 5
Query Match: 44.23% Indels: 0
DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-822-268A-3 (1-48667)
QY 2 GlycerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeu 19
      |||||
Db 15206 GGGAGTCTCTCAGTGTGCTCAAGTCGATCTCAAAATTCCTGGGCTCAAGCGATC 15153

RESULT 9
US-09-864-761-32659/c
; Sequence 32659, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/822,268A
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: 60/211,387
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: NOT YET ASSIGNED
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 48667
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018552.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
; OTHER INFORMATION: NT HIT: D21877.1, EVALUE 1.70e-01
; OTHER INFORMATION: EST_HUMAN HIT: AW592568.1, EVALUE 8.00e-54
; OTHER INFORMATION: SWISSPROT HIT: Q91062, EVALUE 1.70e+00
US-09-864-761-32659

Alignment Scores:
Pred. No.: 8.13 Length: 284
Score: 45.00 Matches: 9
Percent Similarity: 76.92% Conservative: 1
Best Local Similarity: 69.23% Mismatches: 3
Query Match: 43.27% Indels: 0
DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-864-761-32659 (1-284)
QY 4 LeuSerThrPhePheArgLeuPheAsnArgSerPheThr 16
      ::::|
Db 177 GTGAGCACCTCTTCGCGTCTTCCCGCTTCCCGCTCTTGACC 139

RESULT 10
US-09-917-800A-276/c
; Sequence 276, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castile, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
```

; NUMBER OF SEQ ID NOS: 1740
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 276
 ; LENGTH: 333
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AA899635
 ; US-09-917-800A-276

Alignment Scores:
 Pred. No.: 9.87 Length: 333
 Score: 45.00 Matches: 9
 Percent Similarity: 90.91% Conservative: 1
 Best Local Similarity: 81.82% Mismatches: 1
 Query Match: 43.27% Indels: 0
 DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-917-800A-276 (1-333)

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPhe 11
 Db 210 ACGGATCTTATCCAGTTCCTTCGCTTTC 178

RESULT 11

US-09-728-446-245
 ; Sequence 245, Application US/09728446
 ; Patent No. US20020081668A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Friedrich, Glenn
 ; APPLICANT: Zambrowicz, Brian
 ; APPLICANT: Sands, Arthur T.
 ; TITLE OF INVENTION: No. US20020081668A1 Murine Polynucleotide Sequences
 ; FILE REFERENCE: LEX-0101-USA
 ; CURRENT APPLICATION NUMBER: US/09728,446
 ; PRIOR FILING DATE: 2000-11-30
 ; PRIOR FILING DATE: 1999-12-01
 ; NUMBER OF SEQ ID NOS: 1461
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 245
 ; LENGTH: 430
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-09-728-446-245

Alignment Scores:
 Pred. No.: 13.5 Length: 430
 Score: 45.00 Matches: 11
 Percent Similarity: 66.67% Conservative: 3
 Best Local Similarity: 52.38% Mismatches: 5
 Query Match: 43.27% Indels: 2
 DB: 10 Gaps: 1

US-09-833-017B-4 (1-21) x US-09-728-446-245 (1-430)

Qy 2 GlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThr-----GlnAlaLeu 19
 Db 100 GCGGACCTGGACACTTTTCTCTCTTGTAGTCTCCCTGCTCAGCTGACAGACCTT 159

Qy 20 Gly 20
 Db 160 GGC 162

RESULT 12

US-09-733-387-11/c
 ; Sequence 11, Application US/09733387
 ; Patent No. US20020103359A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Donoho, Gregory
 ; APPLICANT: Scoville, John
 ; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Friedrich, Glenn
 ; APPLICANT: Zambrowicz, Brian
 ; APPLICANT: Sands, Arthur T.
 ; TITLE OF INVENTION: No. US20020103359A1 Human Membrane Proteins and
 ; FILE REFERENCE: LEX-0104-USA
 ; CURRENT APPLICATION NUMBER: US/09733,387
 ; PRIOR FILING DATE: 2000-12-07
 ; PRIOR FILING DATE: 1999-12-07
 ; NUMBER OF SEQ ID NOS: 53
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 11
 ; LENGTH: 489
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-733-387-11

Alignment Scores:
 Pred. No.: 15.8 Length: 489
 Score: 45.00 Matches: 9
 Percent Similarity: 76.92% Conservative: 1
 Best Local Similarity: 69.23% Mismatches: 3
 Query Match: 43.27% Indels: 0
 DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-733-387-11 (1-489)

Qy 4 LeuSerThrPhePheArgLeuPheAsnArgSerPheThr 16
 Db 395 GTGAGCACCTCTCCGGTCTTCCCCCGCTCCTTGACC 357

RESULT 13

US-09-974-300-6233/c
 ; Sequence 6233, Application US/09974300
 ; Patent No. US20020146721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berk, Randy M.
 ; APPLICANT: Clausen, Ib Groth
 ; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 ; FILE REFERENCE: 10085.500-US
 ; CURRENT APPLICATION NUMBER: US/09974,300
 ; PRIOR FILING DATE: 2001-10-05
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR FILING DATE: 2000-10-06
 ; NUMBER OF SEQ ID NOS: 8481
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6233
 ; LENGTH: 525
 ; TYPE: DNA
 ; ORGANISM: Bacillus clausii
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(525)
 ; OTHER INFORMATION: n = A,T,C or G
 ; US-09-974-300-6233

Alignment Scores:
 Pred. No.: 17.2 Length: 525
 Score: 45.00 Matches: 9
 Percent Similarity: 65.00% Conservative: 4
 Best Local Similarity: 45.00% Mismatches: 7
 Query Match: 43.27% Indels: 0
 DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-974-300-6233 (1-525)

Qy 2 GlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
 Db 393 GGTTCCTGGGACGGTTCTTCAAAATGAGTAATCTCCGTATACGGGGCCCTGAATAAG 334

RESULT 14

US-09-733-387-31/c
; Sequence 31, Application US/09733387
; Patent No. US20020103359A1
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: NO. US20020103359A1el Human Membrane Proteins and
; FILE REFERENCE: LEX-0104-USA
; CURRENT APPLICATION NUMBER: US/09/733,387
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,427
; PRIOR FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-387-31

Alignment Scores:
Pred. No.: 17.3 Length: 528
Score: 45.00 Matches: 9
Percent Similarity: 76.92% Conservative: 1
Best Local Similarity: 69.23% Mismatches: 3
Query Match: 43.27% Indels: 0
DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-733-387-31 (1-528)

Qy 4 LeuSerThrPhePheArgLeuPheAsnArgSerPheThr 16
Db 395 GTGAGCACCTTCTCCGGTCTTCCCGCCGCTCCTTGACC 357

RESULT 15

US-09-864-761-16165/c
; Sequence 16165, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 16165
; LENGTH: 584
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018552.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
US-09-864-761-16165

Alignment Scores:
Pred. No.: 19.6 Length: 584
Score: 45.00 Matches: 9
Percent Similarity: 76.92% Conservative: 1
Best Local Similarity: 69.23% Mismatches: 3
Query Match: 43.27% Indels: 0
DB: 10 Gaps: 0

US-09-833-017B-4 (1-21) x US-09-864-761-16165 (1-584)

Qy 4 LeuSerThrPhePheArgLeuPheAsnArgSerPheThr 16
Db 435 GTGAGCACCTTCTCCGGTCTTCCCGCCGCTCCTTGACC 397

Search completed: November 9, 2002, 00:58:50
Job time: 27.6866 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: November 8, 2002, 22:01:57 ; Search time 825.582 Seconds
(without alignments)
411.958 Million cell updates/sec

Title: US-09-833-017b-4
Perfect score: 104
Sequence: 1 SCSLSTFFRLFNRSFTQALGK 21

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n_model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09833017/runat_05112002_105349_4846/app_query.fasta_1.398
-DB=EST -QFMT=tastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09833017@cgn.1.1.899 @runat_05112002_105349_4846 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: gb_gss.*
- 18: em_gss_hum.*
- 19: em_gss_inv.*
- 20: em_gss_pln.*
- 21: em_gss_vrt.*
- 22: em_gss_fun.*
- 23: em_gss_mam.*
- 24: em_gss_mus.*
- 25: em_gss_other.*
- 26: em_gss_pro.*
- 27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
c 1	55	52.9	636	17	BH603995
c 2	55	52.9	774	13	BI157225
c 3	54	51.9	691	13	BJ069595
c 4	52.5	50.5	948	17	CNS06VWS
c 5	51	49.0	230	9	AJ504767
c 6	51	49.0	463	14	BQ622211
c 7	51	49.0	494	9	AI236274
c 8	51	49.0	499	13	BI1714045
c 9	51	49.0	544	17	AZ720285
c 10	51	49.0	564	17	BH570656
c 11	51	49.0	730	17	AG019734
c 12	50	48.1	226	17	BH646762
c 13	50	48.1	341	9	AV207860
c 14	50	48.1	364	9	A1762240
c 15	50	48.1	426	9	AA159360
c 16	50	48.1	518	17	BH548953
c 17	50	48.1	577	17	BH708571
c 18	50	48.1	584	17	AQ156697
c 19	50	48.1	603	17	BH492663
c 20	50	48.1	641	17	BH429828
c 21	50	48.1	647	17	AZ639851
c 22	50	48.1	675	17	BH418560
c 23	50	48.1	678	17	AG084609
c 24	50	48.1	710	17	BH438591
c 25	50	48.1	715	17	BH578695
c 26	50	48.1	760	17	BH553018
c 27	50	48.1	775	17	BH521366
c 28	50	48.1	790	17	BH429213
c 29	50	48.1	800	17	BH696657
c 30	50	48.1	817	17	BH605581
c 31	50	48.1	818	17	BH461814
c 32	50	48.1	832	17	BH445663
c 33	50	48.1	835	17	BH587383
c 34	50	48.1	836	17	BH555476
c 35	50	48.1	865	13	BI833209
c 36	50	48.1	939	14	BQ890549
c 37	50	48.1	973	13	BM450250
c 38	49.5	47.6	591	13	BM230202
c 39	49.5	47.6	764	12	BG173055
c 40	49	47.1	235	10	AV979344
c 41	49	47.1	313	17	B23342
c 42	49	47.1	351	10	AW296496
c 43	49	47.1	360	17	AZ467844
c 44	49	47.1	417	10	BB778886
c 45	49	47.1	423	10	AV750133

ALIGNMENTS

RESULT 1
BH603995/c
LOCUS
DEFINITION BH603995 BOGQ Brassica oleracea genomic clone BOGQ193, DNA
sequence.
ACCESSION BH603995
VERSION BH603995.1 GI:17856441
KEYWORDS GSS.
SOURCE Brassica oleracea.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 636)
AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.

ORGANISM Conidiobolus coronatus
Eukaryota; Fungi; zygomycota; Zygomycetes; Entomophthorales;
Ancylistaceae; Conidiobolus.
REFERENCE 1 (bases 1 to 463)
AUTHORS Freimoser, F.M., Screen, S., Hu, G. and St. Leger, R.J.
TITLE EST analysis of genes expressed by the zygomycete pathogen Conidiobolus coronatus during optimized secretion of proteins Unpublished (2002)
JOURNAL
COMMENT Contact: Freimoser F. M.
Department of Entomology
University of Maryland
4112 Plant Sciences Building, College Park, MD 20742, USA
Tel: 301 405 16 13
Fax: 301 314 92 90
Email: ff34@umail.umd.edu.
FEATURES
Location/Qualifiers
1..463
/organism="Conidiobolus coronatus"
/strain="ARSEF 512"
/db_xref="taxon:34488"
/clone_lib="Conidiobolus cornatus ARSEF 512"
/note="Vector: Unizap; Conidiobolus coronatus was grown in minimal medium supplemented with Manduca sexta cuticle and peptone for 18 hours. A cDNA library was constructed in the unidirectional Lambda vector Unizap."
BASE COUNT 157 a 94 c 72 g 134 t 6 others
ORIGIN
Alignment Scores:
Pred. No.: 125 Length: 463
Score: 51.00 Matches: 10
Percent Similarity: 72.22% Conservative: 3
Best Local Similarity: 55.56% Mismatches: 5
Query Match: 49.04% Indels: 0
DB: 14 Gaps: 0
US-09-833-017B-4 (1-21) x BQ622211 (1-463)
QY 3 SerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
||||| ||| |||||||:|||||:|||||
Db 233 TCAATTGAGNTGTTTCATGTTATTTTCAAGATCTTTTACAGAGCGGTGGGT 180
RESULT 7
LOCUS AI236274
DEFINITION EST232836 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone ROVD54 3' end, mRNA sequence.
ACCESSION AI236274.1 GI:3829780
VERSION
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS Lee, N.H., Glodek, A., Chandra, I., Mason, T.M., Quackenbush, J., Kerlavage, A.R. and Adams, M.D.
TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat Gene Index
JOURNAL Unpublished (1998)
COMMENT Other_ESTS: TC61188
Contact: Lee, NH
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@tigr.org
Seq primer: M13-21.
FEATURES
Location/Qualifiers
1..494
/organism="Rattus sp."
/db_xref="ATCC (inhost):2041604"
/db_xref="taxon:10118"
/clone_lib="ROVD54"
/note="Organ: ovary; Vector: pT7T3pac; Site_1: EcoRI; Site_2: NotI"
BASE COUNT 150 a 114 c 113 g 117 t
ORIGIN
Alignment Scores:
Pred. No.: 135 Length: 494
Score: 51.00 Matches: 13
Percent Similarity: 75.00% Conservative: 2
Best Local Similarity: 65.00% Mismatches: 5
Query Match: 49.04% Indels: 0
DB: 9 Gaps: 0
US-09-833-017B-4 (1-21) x AI236274 (1-494)
QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
||||| |||||||:|||||:|||||
Db 315 TCCGGATCTCTCATCATATATCCAGACTTCCCAACCTATCAAAAACCTGGAGCACTTGGT 374
RESULT 8
LOCUS BI714045/c
DEFINITION BI714045 499 bp mRNA linear EST 19-SEP-2001
ie31e06.xl1 Kaestner ngn3 wt Mus musculus cDNA 3' similar to TR:Q921Q8 Q921Q8 HYPOTHETICAL 94.9 KD PROTEIN. ;, mRNA sequence.
ACCESSION BI714045.1 GI:15689740
VERSION
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A., Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y. and Bowers, Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohph.harvard.edu
Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611, 2000) Library was constructed by Catherine Lee DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Marie Searce (mscearce@mail.med.upenn.edu)
Possible reversed clone: similarity on wrong strand
Seq primer: -400P from Gibco
High quality sequence stop: 427.
FEATURES
Location/Qualifiers
1..499
/organism="Mus musculus"
/strain="129/Sv x CD1"
/db_xref="taxon:10090"
/clone_lib="Kaestner ngn3 wt"
/dev_stage="p.c. 14.5"
/lab_host="E. coli-DH12S (GIBCO)"
/note="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1: Not I; Site_2: Sal I; The library was prepared by Catherine S. Lee and has not been published. The pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611, 2000). The cDNA's were prepared with an oligo containing a NotI site, and SalI linkers were added to the ends. The

inserts were cut with NotI before being cloned into the NotI-Sali sites in the vectors. This is one of two libraries, ngn3 wt and ngn3 -/- . The wt library is in pSPORT1, T7 promoter is 5'.

BASE COUNT 144 a 104 c 127 g 124 t

ORIGIN

Alignment Scores:

Pred. No.: 136 Length: 499

Score: 51.00 Matches: 12

Percent Similarity: 70.59% Conservative: 0

Best Local Similarity: 70.59% Mismatches: 5

Query Match: 49.04% Indels: 0

DB: 13 Gaps: 0

US-09-833-017B-4 (1-21) x B1714045 (1-499)

QY 5 SerThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
||||| ||||||| ||| ||||||| |||||||

Db 59 AGTACAGTGTTCAGACTGGAAATAGCTCTTCACATCTTTTGGGAAA 9

RESULT 9

AZ720285 544 bp DNA linear GSS 24-JAN-2001

LOCUS RPCI-24-9201.TJ RPCI-24 Mus musculus genomic clone RPCI-24-9201,
DNA sequence.

ACCESSION AZ720285

VERSION AZ720285.1 GI:12461826

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsedaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P., and Fraser,C.M.

TITLE Mouse BAC End Sequences from Library RPCI-24

JOURNAL Unpublished (1999)

COMMENT Other_GSSs: RPCI-24-9201.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end

page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 92 row: 0 column: 1

Seq primer: SP6

Class: BAC ends.

FEATURES Location/Qualifiers

1..544

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-9201"

/clone_lib="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/note="Vector: PTARBAC1; Site_1: BamHI; Site_2: BamHI;

RPCI-24 Mouse BAC Library produced by Pieter de Jong. The

library was cloned in the PTARBAC1 cloning vector at the

BamHI sites using MboI partially digested male C57BL/6J

DNA."

BASE COUNT 175 a 114 c 129 g 126 t

ORIGIN

Alignment Scores:

Pred. No.: 151 Length: 544

Score: 51.00 Matches: 9

Percent Similarity: 80.00% Conservative: 3

Best Local Similarity: 60.00% Mismatches: 3

Query Match: 49.04% Indels: 0

DB: 17 Gaps: 0

US-09-833-017B-4 (1-21) x AZ720285 (1-544)

QY 6 ThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
||| ||||||| ||| ||||||| ||| |||||||

Db 19 ACAATATTTCGCTTGTAAATAGACTTTCAATCAGAGTCTAGGC 63

RESULT 10

BH570656/c

LOCUS BH570656 564 bp DNA linear GSS 14-DEC-2001

DEFINITION BOHMO30TR BOHM Brassica oleracea genomic clone BOHMO30, DNA
sequence.

ACCESSION BH570656

VERSION BH570656.1 GI:17822495

KEYWORDS GSS.

SOURCE Brassica oleracea.

ORGANISM Brassica oleracea

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.

AUTHORS Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.

TITLE Whole genome shotgun sequencing of Brassica oleracea

JOURNAL Unpublished (2001)

COMMENT Other_GSSs: BOHMO30TF

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES Location/Qualifiers

1..564

/organism="Brassica oleracea"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOHMO30"

/clone_lib="BOHM"

/note="Vector: pHOS1; Site_1: BstXI; 2-3 kb sheared

genomic DNA inserted into pHOS1 using BstXI linkers"

BASE COUNT 157 a 163 c 120 g 124 t

ORIGIN

Alignment Scores:

Pred. No.: 157 Length: 564

Score: 51.00 Matches: 11

Percent Similarity: 73.68% Conservative: 3

Best Local Similarity: 57.89% Mismatches: 5

Query Match: 49.04% Indels: 0

DB: 17 Gaps: 0

US-09-833-017B-4 (1-21) x BH570656 (1-564)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeu 19
||||| ||||||| ||||||| ||||||| |||||||

Db 323 TCAGGTGAGCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 267

RESULT 11

AG019734

LOCUS AG019734 730 bp DNA linear GSS 16-OCT-1999

DEFINITION Homo sapiens genomic DNA, 21q region, clone: B2289H10 N056(RP),
genomic survey sequence.

ACCESSION AG019734

VERSION AG019734.1 GI:6045678

```

KEYWORDS
SOURCE      Homo sapiens DNA, clone:B2289H10 N056(RP).
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
JOURNAL    Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
REFERENCE   Homo sapiens genomic DNA, chromosome 21q
AUTHORS    Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
JOURNAL    Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL    Submitted (16-OCT-1999) Masahira Hattori, RIKEN Genomic Sciences
           Center(GSC) c/o Kitasato University; 1-15-1 Kitasato, Sagamihara
           228-8555, Japan (E-mail:hattori@gsc.ims.u-tokyo.ac.jp,
           Tel:042-778-9923, Fax:042-778-9924)
FEATURES   Location/Qualifiers
           source
             1..730
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="21"
               /map="21q"
BASE COUNT 192 a 141 c 127 g 270 t
ORIGIN
Alignment Scores:
Pred. No.: 211 Length: 730
Score: 51.00 Matches: 11
Percent Similarity: 70.00% Conservative: 3
Best Local Similarity: 55.00% Mismatches: 6
Query Match: 49.04% Indels: 0
DB: 17 Gaps: 0

US-09-833-017B-4 (1-21) x AG019734 (1-730)

Qy 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
Db 370 ACAGGATCCATTCCACATGAATCCTTATTTCATATTCTTGGGTCACAGGATTAGGA 429

RESULT 12
BH646762/c
LOCUS      BH646762
DEFINITION BH646762.1 GI:18704175
ACCESSION BH646762
VERSION    BH646762.1
KEYWORDS   Brassica oleracea.
SOURCE     Brassica oleracea.
ORGANISM   Brassica oleracea
REFERENCE   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
JOURNAL    Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
TITLE      1 (bases 1 to 226)
AUTHORS    Town,C.D., Van Aken,S., Utterback,T. and Fraser,C.M.
JOURNAL    Whole genome shotgun sequencing of Brassica oleracea
COMMENT    Unpublished (2001)
           Contact: Chris Town
           TIGR
           9712 Medical Center Drive, Rockville, MD 20850, USA.
           Tel: 301-838-3523
           Fax: 301-838-0208
           Email: cdtown@tigr.org
           DNA is from a doubled haploid provided by Tom Osborn.
           Seq primer: TF
           Class: sheared ends.
FEATURES   Location/Qualifiers
           source
             1..226
               /organism="Brassica oleracea"
               /strain="T01000DH3"
               /db_xref="taxon:3712"

```

```

/clone="BOMFC55"
/clone_lib="BQ_2_3_KB"
/note="Vector: pHOS1; Site1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"
BASE COUNT 47 a 50 c 54 g 75 t
ORIGIN
Alignment Scores:
Pred. No.: 80.7 Length: 226
Score: 50.00 Matches: 10
Percent Similarity: 76.47% Conservative: 3
Best Local Similarity: 58.82% Mismatches: 4
Query Match: 48.08% Indels: 0
DB: 17 Gaps: 0

US-09-833-017B-4 (1-21) x BH646762 (1-226)

Qy 5 SerThrPheArgLeuPheAsnArgSerPheThrGlnAlaLeuGlyLys 21
Db 64 GCAACTTACCAACGACTCGTGAACGCTATGTTCTCTCAACAACCTCGGTRAA 14

RESULT 13
AV207860
LOCUS      AV207860
DEFINITION AV207860 RIKEN full-length enriched, adult male testis Mus musculus
           cDNA clone 1700099M17 3', mRNA sequence.
ACCESSION AV207860
VERSION    AV207860.1 GI:6148713
KEYWORDS   EST.
SOURCE     house mouse.
ORGANISM   Mus musculus
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
           1 (bases 1 to 341)
           Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
           Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
           Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I., Kai
           ,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
           Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
           Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibata,K., Shibata
           ,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugahara,Y., Suzuki,H.,
           Suzuki,H., Takahashi,F., Tateo,M., Tomimaga,N., Tsunoda,Y.,
           Watahiki,A., Watanabe,S., Yamamura,T., Yasunishi,A., Yokota,T.,
           Yoshiki,N., Yashino,M., Muramatsu,M. and Hayashizaki,Y.
           RIKEN Mouse ESTs (Konno,H., et al. 1999)
           Unpublished (1999)
           Contact: Yoshihide Hayashizaki
           Laboratory for Genome Exploration Research Group, RIKEN Genomic
           Sciences Center(GSC), Yokohama Institute
           The Institute of Physical and Chemical Research (RIKEN)
           1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
           Tel: 81-45-503-9222
           Fax: 81-45-503-9216
           Email: genome-res@sc.riken.go.jp,
           URL:http://genome.gsc.riken.go.jp/
           Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
           Matsunura,S., Carninci,P., Muramatsu,M., Okazaki,Y. and Hayashizaki
           ,Y.
           Transcriptional sequencing: A method for DNA sequencing using RNA
           polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3453-3460 (1998)
           Itoh,M., Kitsuai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
           Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki
           ,Y. and Hayashizaki,Y.
           Automated filtration-based high-throughput plasmid preparation
           system. Genome Res. 9 (5), 463-470 (1999)
           Carninci,P. and Hayashizaki,Y.
           High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
           19-44 (1999)
           Please visit our web site (http://genome.rtc.riken.go.jp) for
           further details.
FEATURES   Location/Qualifiers
           source
             1..341
               /organism="Mus musculus"

```


ORIGIN

Alignment Scores:
Pred. No.: 167 Length: 426
Score: 50.00 Matches: 11
Percent Similarity: 66.67% Conservative: 3
Best Local Similarity: 52.38% Mismatches: 7
Query Match: 48.08% Indels: 0
DB: 9 Gaps: 0

US-09-833-017B-4 (1-21) x AAL59360 (1-426)

QY 1 SerGlySerLeuSerThrPhePheArgLeuPheAsnArgSerPheThrGlnAlaLeuGly 20
||| :::::::::::|||||::|
Db 23 TCAGTAGCATTATCCAGCTTTTCCGTTTGTCCAGAAATAGTCTTACTGGCCTTGGT 82

QY 21 Lys 21
|||

Db 83 AAA 85

Search completed: November 8, 2002, 23:39:49
Job time : 830.582 secs